



مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الثانية
(٢٩)

أساسيات استرجاع المعلومات (نظم استرجاع المعلومات)

تأليف

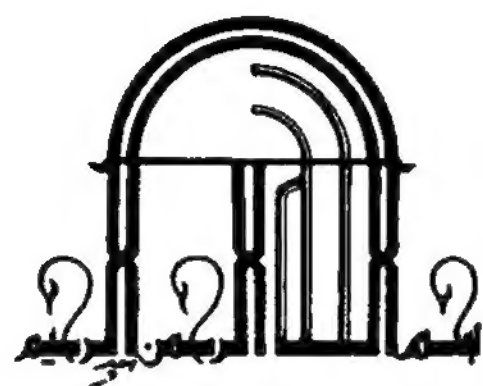
ف. و. لانكستر و أ. ج. وورنر

ترجمة

د. حشمت قاسم

الرياض

١٤١٨هـ / ١٩٩٧م



مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الثانية (٢٩)

تهتم هذه السلسلة بنشر الدراسات والبحوث في
إطار علم المكتبات والمعلومات بشكل عام

أساسيات استرجاع المعلومات

(نظم استرجاع المعلومات)

تأليف

ف. و. لانكستر و أ.ج. وورنر

ترجمة

الدكتور / حشمت قاسم

الطبعة الثالثة، مزيده ومنقحه

مكتبة الملك فهد الوطنيه

١٤١٨هـ - ١٩٩٧م

ح) مكتبة الملك فهد الوطنية ، ١٤١٨ هـ

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

لانكستر ، فردرك ولغرد

أساسيات إسترجاع المعلومات/ فردرك ولغرد لانكستر ، آمي ج. وورنر ، ترجمة حشمت قاسم .- الرياض.

٤٥٤ ص ، ٢٤ سم .- (السلسلة الثانية ، ٢٩)

ردمك ٩٩٦٠-٠٠-١٠٦-٧

ردمك ١٣١٩-٢٩٨١

١- إختزان وإسترجاع المعلومات أ - وورنر ، آمي ج. (م. مشارك)

ب- قاسم ، حشمت (مترجم) ج- العنوان د - السلسلة

١٨/٢٩٨٥

ديوي ٢٩٠٧

رقم الإيداع : ١٨/٢٩٨٥

ردمك : ٩٩٦٠-٠٠-١٠٦-٧

ردمك : ١٣١٩-٢٩٨١

جميع حقوق الطبع محفوظة ، غير مسموح بطبع أي جزء من أجزاء هذا الكتاب ، أو اختراعه في أي نظام لاختزان المعلومات وإسترجاعها ، أو نقله على أية هيئة أو بأية وسيلة سواء كانت إلكترونية أو شرائط ممغنطة أو ميكانيكية ، أو استنساخا ، أو تسجيلا ، أو غيرها إلا في حالات الاقتباس المحدودة بغرض الدراسة مع رجوب ذكر المصدر .

ص ب : ٧٥٧٢

الرياض : ١١٤٧٢ المملكة العربية السعودية

هاتف : ٤٦٢٤٨٨٨

فاكس - ناسوخ : ٤٦٤٥٣٤١

هذه ترجمة كاملة لكتاب :

Lancaster, F. Wilfrid and Amy J. Warner.

Information Retrieval Today. Revised, retitled

and expanded edition of "Information retrieval systems..."

Arlington, Virginia: Information Resources, prees, 1993.

طبعت هذه الترجمة باتفاق مع الناشر

قائمة المحتويات

| الصفحة | الفقرة |
|--------|--|
| ١١ | مقدمة المترجم |
| ١٣ | تصدير |
| ١٥ | الفصل الأول - بعض أساسيات استرجاع المعلومات . |
| ١٥ | دورة تداول المعلومات |
| ١٨ | دور مرافق المعلومات |
| ٢٣ | نظم استرجاع المعلومات |
| ٢٣ | التحليل الموضوعي للوثائق |
| ٣٠ | مشكلات استرجاع المعلومات |
| ٣٣ | مكونات نظم استرجاع المعلومات - |
| ٣٤ | المكتبة كنظام لاسترجاع المعلومات |
| ٤٣ | الفصل الثاني - صناعة مراصد البيانات |
| ٤٤ | نظم العمل على دفعات |
| ٤٦ | نظم الخط المباشر |
| ٥١ | الاتجاهات السائدة في الصناعة |
| ٥٤ | منتجات المعلومات وخدماتها |
| ٦٢ | متعهد ومراصد البيانات . |
| ٧٣ | الفصل الثالث - الاسترجاع الموضوعي ؛ مشكلاته ومعايير الأداء |
| ٧٣ | مقاييس كفاءة الاسترجاع |
| ٧٦ | العوامل المؤثرة في كفاءة نظم الاسترجاع |
| ٧٨ | الصلاحية والاتصال بالموضوع |
| ٨٣ | الاتصال بالموضوع |
| ٨٥ | الصلاحية . |
| ١٠١ | الفصل الرابع - تواصل المستفيد والوسيط |
| ١٠١ | الاحتياجات في مقابل الطلبات |

| | |
|-----|---|
| ١٠٩ | نوعية طلبات المستفيدين |
| ١١٢ | طرق تفاعل المستفيد والوسيط |
| ١١٦ | نماذج طلب البحث |
| ١١٩ | الفصل الخامس - التحليل الموضوعي |
| ١٢١ | الكشف بواسطة البشر |
| ١٢٧ | الكشف الجزئي وكشف النصوص كاملة |
| ١٣٠ | الخلاصة |
| ١٣٣ | الفصل السادس - اللغة في الاسترجاع |
| ١٣٤ | خصائص اللغة واللغة المقيدة |
| ١٣٦ | الاشتراك اللفظي |
| ١٣٧ | البحث الشامل |
| ١٣٧ | وظائف اللغة المقيدة |
| ١٣٨ | اختيار المصطلحات |
| ١٤٧ | اللغة المقيدة |
| ١٥٣ | خطط التصنيف |
| ١٥٤ | مظاهر التطوير في التصميم |
| ١٥٤ | اللغة الطبيعية في استرجاع المعلومات |
| ١٥٩ | الفصل السابع - اختيار مرادف البيانات وتقييمها |
| ١٦٠ | المصادر المطبوعة |
| ١٦٢ | كشافات مرادف البيانات |
| ١٦٨ | الوسائل الآلية - الدراسات المتصلة |
| ١٦٩ | نوعية مرادف البيانات |
| ١٧٠ | التغطية |
| ١٨٢ | القابلية للتنبؤ |
| ١٨٣ | الفورية |
| ١٨٣ | آفاق المستقبل |
| ١٨٧ | الفصل الثامن - البحث في مرادف البيانات |
| ١٨٨ | منطق البحث |
| ١٩٢ | استراتيجية البحث |

| | |
|-----|---|
| ١٩٥ | توسعة استراتيجيات البحث وتضييقها |
| ٢٠١ | بعض عينات عمليات البحث |
| ٢١٠ | بحث النصوص الحرة |
| ٢١٥ | البحث بالمصطلحات الموزونة |
| ٢١٨ | البحث الجزئي - غربلة المخرجات |
| ٢١٩ | خصائص مسئول البحث |
| ٢٢٢ | العوامل المؤثرة في عملية بحث بعينها |
| ٢٢٥ | الفصل التاسع - معايير التقييم وإجراءاته |
| ٢٢٥ | المعايير المستخدمة في تقييم نظم استرجاع المعلومات |
| ٢٢٨ | مستويات التقييم |
| ٢٢٩ | الاستدعاء والتحقيق |
| ٢٣٧ | المقاييس الأخرى لتقييم عمليات البحث |
| ٢٤٠ | إجراء التقييم |
| ٢٤٨ | الحصول على أرقام الأداء |
| ٢٥٧ | تفسير النتائج |
| ٢٦٢ | دراسات فعالية التكلفة |
| ٢٧٣ | تحليل عائد التكلفة |
| ٢٨٣ | الفصل العاشر - العوامل المؤثرة في الأداء في استرجاع المعلومات |
| ٢٨٥ | العوامل الخاصة بالنظام في استرجاع المعلومات ، مرصد البيانات |
| ٢٨٦ | التغطية |
| ٢٨٩ | عوامل الوقت ، التكشيف واللغة |
| ٢٩٠ | الجوانب الخاصة بفعالية تكلفة مرصد البيانات |
| ٢٩٥ | لغة التكشيف ، التخصيص |
| ٣٠٢ | الربط |
| ٣٠٣ | البناء |
| ٣٠٦ | التكشيف |
| ٣١١ | التخصيص |
| ٣١٣ | الدقة |
| ٣١٤ | جوانب فعالية التكلفة في التكشيف |

| | |
|-----|--|
| ٣١٩ | استراتيجية البحث |
| ٣٢٥ | جوانب فعالية التكلفة في عمليات البحث |
| ٣٢٧ | عمليات المواءمة في استرجاع المعلومات |
| ٣٣٢ | العوامل البشرية في التكشيف والبحث |
| ٣٣٣ | دراسات الاطراد |
| ٣٣٦ | دراسات الكفاءة |
| ٣٤١ | الفصل الحادي عشر - الطرق الآلية لاسترجاع المعلومات |
| ٣٤٧ | الاساليب اللغوية لاسترجاع المعلومات |
| ٣٥٢ | الطرق الإحصائية لاسترجاع المعلومات |
| ٣٥٨ | فئات الطرق الآلية لاسترجاع المعلومات |
| ٣٦٠ | بناء المكانز آلياً |
| ٣٦٥ | التكشيف الآلي |
| ٣٧١ | صياغة الاستفسارات آلياً |
| ٣٧٧ | الاستخلاص الآلي |
| ٣٨١ | الطرق الآلية المستخدمة في الفهارس المتاحة على الخط |
| ٣٧٣ | تقييم الطرق الآلية |
| ٣٨٧ | الفصل الثاني عشر - الاتجاهات والتطورات المستقبلية المحتملة |
| ٣٨٩ | نمو مرصد البيانات وتقنيات المعلومات |
| ٣٩٣ | استرجاع المعلومات وأطر استخدام الحاسبات |
| ٣٩٥ | التحول من المطبوعات إلى الوثائق الإلكترونية |
| ٤٠٠ | بيئة المكتبات ومراكز المعلومات |
| ٤٠٦ | آراء أخرى حول المستقبل |
| ٤١٣ | الذكاء الاصطناعي والنظم الخبيرة |
| ٤٢١ | قائمة المراجع |
| ٤٤٥ | معجم المصطلحات |

مقدمة المترجم

البشر إزاء التطورات الحديثة صنفان ؛ صنف يقف عند حدود الانبهار السلبي ، وصنف قادر على تتبع الجذور ورصد المراحل واستكشاف العوامل والأسباب ، ومن ثم استشراف آفاق المستقبل بوعي ورؤية ثاقبة . وينتمي فردرك ولفرد لانكستر إلى الصنف الثاني ، حيث لم تبهره التطورات التقنية التي شهدتها مجال المعلومات خلال العقدين الأخيرين ، ويضع هذه التطورات في سياقها المناسب ، في تناغم مع التطورات العلمية الأساسية في المجالات المتصلة بقضايا المعالجة الموضوعية لأوعية المعلومات ، وتصميم نظم الاسترجاع ، وسبل التعامل مع هذه النظم ، ومعايير تقييمها ، واتجاهات نموها وتطورها ، وهي مجالات تمتد ، كما نعلم ، عبر نطاق عريض من التخصصات ، يشمل من بين ما يشمل ، علم اللغة مع التركيز بوجه خاص على علم اللغة الحاسبي ، وتطبيقات الحاسب الآلي وخاصة في الذكاء الاصطناعي ، وتقنيات النشر والاستنساخ المتصلة بالأوعية الإلكترونية والأسطوانات الضوئية المكتنزة ، وتقنيات الاتصالات بعيدة المدى والمشابكة ، فضلاً عن الإدارة والعلوم السلوكية .

في هذا الإطار المعرفي يعالج ولفرد لانكستر وأمي وورنر الوضع الراهن لمجال استرجاع المعلومات ، في هذا الكتاب الذي يمثل امتداداً لكتاب لانكستر الأصلي « نظم استرجاع المعلومات » الذي صدر في طبعتين ، الأولى عام ١٩٦٨ م والثانية عام ١٩٧٩ م ، وكان لنا شرف ترجمة الطبعة الثانية إلى العربية عام ١٩٨١ م . وعلى ذلك فإن الكتاب الذي بين أيدينا هو الطبعة الثالثة التي جاءت بعد مضي أربعة عشر عاماً من صدور الطبعة الثانية ، ونتشرف بتقديم ترجمتها العربية بعد خمسة عشر عاماً من ترجمة

الطبعة الثانية . وقد شهدت هذه السنوات تطورات جوهرية في مراصد البيانات ونظم استرجاع المعلومات ، وتقنيات الحاسبات والاتصالات وقنوات بث المعلومات ، ترصدها هذه الطبعة بوعي يؤكد تواصل الحلقات ، ويبرز تضافر الروافد والمقومات ، ويمهد الطريق لما يلي من خطوات .

وفردرك ولفرد لانكستر غني عن التعريف ، وقد سبق أن عرفنا به في التمهيد للترجمة العربية للطبعة الثانية من هذا الكتاب ، ويكفي القول في هذا السياق إنه من أبرز أساتذة علم المعلومات المعاصرين ، ومن أغزر الثقات إنتاجاً . ويعمل لانكستر الآن أستاذاً للدراسات المتقدمة بالمعهد العالي للمكتبات وعلم المعلومات بجامعة إلينوي بالولايات المتحدة الأمريكية ، ويرأس تحرير واحدة من أعرق الدوريات المتخصصة في مجال المكتبات ، وهي *Library Trends* ، وله ثمانية كتب في علم المعلومات ، حصل خمسة منها على جوائز قومية . ويتمتع لانكستر بخبرة عالمية واسعة ، اكتسبها من مشاركاته المكثفة بالجهود الاستشارية في عدد كبير من الدول .

وفضلاً عن استيعاب التطورات التي تلت صدور الطبعة الثانية ، فإن هذه الطبعة الثالثة تختلف في مبنائها ومحتواها . فبينما كان الكتاب في طبعته الثانية يضم خمسة وعشرين فصلاً بالإضافة إلى ثلاثة ملاحق ، ينتظم محتوى الطبعة الثالثة اثنا عشر فصلاً . وبينما كان تقييم نظم استرجاع المعلومات موزعاً على ثمانية فصول في الطبعة الثانية ، تركّز هذا الموضوع في فصلين اثنين فقط في الطبعة الثالثة . وبينما كانت لغويات نظم استرجاع المعلومات مشتتة في ثانيا عدد من فصول الطبعة الثانية ، تركّز الاهتمام بهذا الموضوع في فصل مستقل في الطبعة الثالثة . وهذه مجرد أمثلة لمظاهر الاختلاف بين الطبعتين ، ولا ننسى في هذا الصدد وجود مؤلف مشارك في الطبعة الثالثة .

وإذ نقدم هذه الترجمة إلى القارئ العربي ، نرجو أن نكون قد وفقنا في نقل رسالة المصدر كاملة دون تحريف أو تشويه . وما توفيقني إلا بالله عليه توكلت وإليه أنيب ، ونسأله سبحانه وتعالى أن يهدينا سواء السبيل .

الدكتور حشمت قاسم

القاهرة في رمضان ١٤١٦ هـ

فبراير ١٩٩٦ م

تصدير

شهد مجال استرجاع المعلومات الكثير من الأحداث والتطورات منذ صدور الطبعة الثانية من هذا الكتاب عام ١٩٧٩م (*) . فقد تزايدت وبشكل ملحوظ أعداد مراصد البيانات الإلكترونية ، وشبكات الاتصالات بعيدة المدى ، التي يمكن عن طريقها التعامل مع هذه المراصد ، كما انتشرت مراصد البيانات المسجلة على الأسطوانات الضوئية المكتنزة CD-ROM في المكتبات على اختلاف أنواعها وتفاوت أحجامها ، كما تزايد أيضاً عدد مراصد البيانات التي تضم النصوص الكاملة لأوعية المعلومات ، وها هي الفهارس المتاحة للجمهور على الخط المباشر (OPACS) تحل إلى حد بعيد ، محل الفهارس البطاقية .

وقد أدت هذه التطورات المتنوعة إلى تزايد الاهتمام في الوقت الراهن بالاسترجاع الموضوعي ، عما كان عليه من قبل . وما زال هناك المزيد من جهود البحث الجارية في هذا الميدان . ويبدو أن كثيرين قد أقنعوا أنفسهم بقدرة التقنيات وحدها على حل مشكلات الاسترجاع الموضوعي ، إلا أن هذه المشكلات ما زالت في انتظار الحل ، على الرغم من بعض غزوات النظم الخبيرة المبنية على المعرفة ، وغير ذلك مما يقدمه الحاسب الآلي من مساعدات في عملية التكشيف . كذلك تحمل أساليب التناول اللغوية بين طياتها ما يبشر بالأمل . وفضلاً عن هذين المجالين (النظم الخبيرة وأساليب التناول اللغوية) تتناول هذه الطبعة الاسترجاع الموضوعي الترابطي ، والنصوص الفائقة Hypertext ، والاسترجاع غير الموضوعي ، وغير ذلك من القضايا الراهنة للمعلومات .

(*) ولغرد لانكستر . نظم استرجاع المعلومات ، ترجمة حشمت قاسم . القاهرة ، مكتبة غريب ، ١٩٨١ م .

وتواصل هذه الطبعة الثالثة لكتاب Char- Information Retrieval Systems : Characteristics, Testing and Evaluation الذي صدر لأول مرة عام ١٩٦٨م، التأكيد على أن مشكلات استرجاع المعلومات الحقيقية ، هي المشكلات الفكرية الخاصة بالتحليل الموضوعي والتعبير عن نتائج هذا التحليل ، وأنه ما من حل لهذه المشكلات اعتماداً على التقنيات دون سواها .

وكما هو الحال في الطبعتين السابقتين ، فقد قصد بهذا الكتاب ، في الأساس أن يكون نصاً يعتمد عليه في المقررات الدراسية لمجال استرجاع المعلومات ، كما يأمل المؤلفان أن تكون له أهميته وقيمه أيضاً بالنسبة للمكتبيين ، ومنتجي مرصدي البيانات ، وغيرهم من الاختصاصيين المهتمين بمشكلات الاسترجاع الموضوعي .

فردريك ولفرد لانكستر

آمي ج . وورنر

الفصل الأول

بعض أساسيات استرجاع المعلومات

دورة تداول المعلومات :

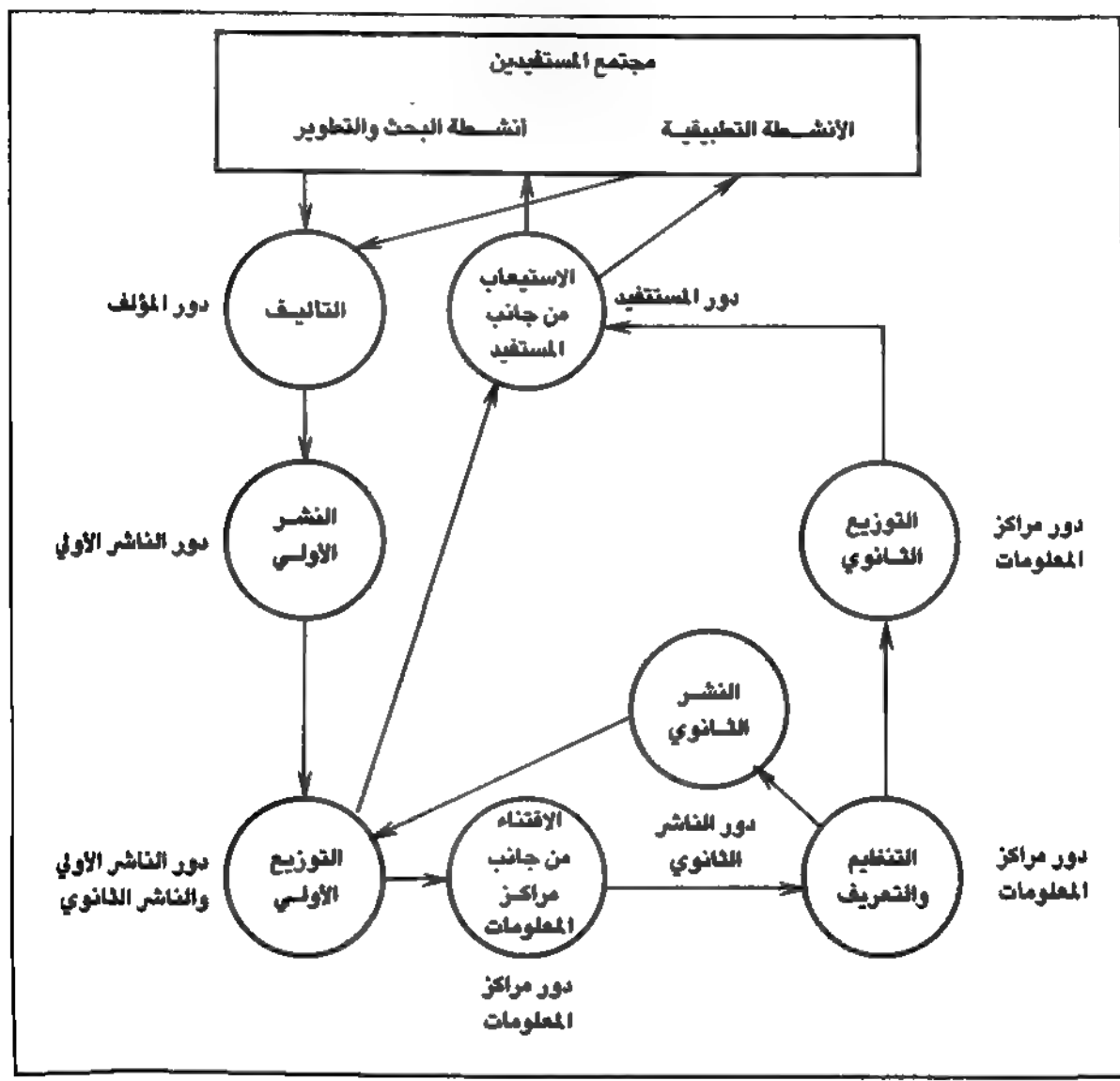
استرجاع المعلومات هو النشاط الأساس فيما تقوم به مراكز المعلومات . وتشمل مراكز المعلومات أو مرافق المعلومات المكتبات ومنتجي مراصد البيانات المنشورة في شكل مطبوع أو في شكل إلكتروني ، وأية نوعية أخرى من المرافق التي تقدم مصادر المعلومات لمجتمع من المستفيدين . ومن الممكن إدراك الوظائف المهمة التي تضطلع بها مراكز المعلومات ، وفهمها على أحسن وجه إذا ما نظرنا إليها في السياق العريض للدورة التي يتم بها تداول المعلومات عبر القنوات الرسمية .

ويصور الشكل رقم (١) العناصر الأساسية للدورة . و « مجتمع المستفيدين » هو هؤلاء الأفراد العاملون في مجال موضوعي معين . وبعض هؤلاء الأفراد يمارسون « أنشطة البحث والتطوير » والبعض الأخرى يمارس أنواعاً متعددة من الأنشطة الأخرى التي يشار إليها إجمالاً « بالأنشطة التطبيقية » . وجميع هؤلاء مستفيدون من المعلومات ، وبعضهم منتج للمعلومات أيضاً ، أي أن هؤلاء الذين يمكن لأنشطتهم أن تروق للآخرين في المجتمع ، يسجلون خبراتهم أو بحوثهم أو آراءهم في نوع ما من التقارير . وهذا النشاط هو « دور المؤلف » في دورة تداول المعلومات . إلا أن التأليف ليس في حد ذاته شكلاً من أشكال الاتصال ؛ فمن الممكن لجهد المؤلف أن يكون ضعيف التأثير أو عديم التأثير كلية في مجتمع المتخصصين ، إلى أن يستنسخ في عدة نسخ ويوزع رسمياً ، أي ينشر . والنشر هو « دور الناشر الأولي » في دورة تداول المعلومات .

ويمكن للمطبوع الأولي أن يكون كتاباً ، أو دورية ، أو تقريراً تقنياً ، أو أطروحة ، أو براءة اختراع ، أو ما شابه ذلك (*) . وتوزع المطبوعات الأولية على مجتمع المستفيدين بطريقتين :

١ - مباشرة عن طريق الاشتراك والشراء من جانب الأفراد .

٢ - بشكل غير مباشر عن طريق الاشتراك والشراء من جانب مراكز المعلومات .



الشكل رقم (١) : دورة تداول المعلومات .

ولمراكز المعلومات أدوار مهمة جداً في دورة تداول المعلومات ؛ فمن خلال سياساتها الخاصة بالاقتناء والاختزان تظل المكتبات أرشيفاً دائماً للإنجازات المهنية ،

(*) يبالغ المؤلفان هنا في التبسيط في مفهوم النشر ؛ فالأطروحات وتقارير البحوث ، على عكس الكتب والدوريات ، ليست من الأوعية التي يتم تداولها بلا قيد ولا شرط . (المترجم) .

ومصدراً آمناً للإفادة من هذه التسجيلات . هذا بالإضافة إلى أن المكتبات وغيرها من مراكز المعلومات تنظم الإنتاج الفكري وتعرف به عن طريق الفهرسة والتصنيف والتكشيف وما يرتبط بذلك من إجراءات . كذلك تقوم مؤسسات التكشيف والاستخلاص وناشرو الوراقيات الوطنية أيضاً بتنظيم الإنتاج الفكري والتعريف به ، حيث تتولى مسئولية نشر وتوزيع « المطبوعات الثانوية » . وبعض هذه المطبوعات الثانوية يتجه مباشرة إلى مجتمع المستفيدين ، وإن كانت الغالبية العظمى منها تتجه نحو المشتركين من المؤسسات (مراكز المعلومات) .

كذلك تضطلع مراكز المعلومات أيضاً بمهام العرض والبت الحيوية في الدورة . وتشمل هذه الأنشطة التي تمثل شكلاً من أشكال التوزيع الثانوي للمطبوعات والمعلومات عن المطبوعات ، إعارة الأوعية أو المواد ، ومختلف أنواع خدمات الإحاطة الجارية والخدمات المرجعية وخدمات البحث في الإنتاج الفكري . ولما رصد البيانات الإلكترونية التي ينتجها ناشرو الأوعية الثانوية ، أهميتها المتزايدة منذ سبعينيات القرن العشرين ، في تقديم مختلف أنواع خدمات المعلومات .

و « الاستيعاب » هو المرحلة الأخيرة في دورة تداول المعلومات . وهذه المرحلة ، التي لا يمكن أن نلمسها ، هي المرحلة التي يتم فيها استيعاب المعلومات من جانب مجتمع المستفيدين . ونميز هنا بين تداول الوثائق وتداول المعلومات ؛ فتداول المعلومات لا يتم فعلاً إلا عندما تُدرس إحدى الوثائق من جانب أحد المستفيدين ، حيث يتم استيعاب محتواها إلى الحد الذي يصبح فيه المستفيد عليماً بهذا المحتوى ، أي تتغير معرفته بالموضوع . ومن الممكن لاستيعاب المعلومات من جانب مجتمع المتخصصين أو المهنيين أن يتم عن طريق التوزيع الأولي أو التوزيع الثانوي . ويرتبط اختلاف مستويات وسرعات الاستيعاب باختلاف الوثائق ، كما يمكن لمحتوى بعضها ألا يستوعب على الإطلاق نظراً لأن أحداً لم يفد منها كلية .

هذا ، وقد تم تصوير عمليات الاتصال الرسمي في شكل دائري نظراً لأنها تتسم بالاستمرارية والتجدد . ويمكن للقراء عن طريق عملية الاستيعاب تحصيل المعلومات التي يمكن أن يفيدوا منها في أبحاثهم وأنشطتهم التطويرية ، التي تؤدي بدورها إلى الكتابة والنشر ، وبذلك تستمر الدائرة .

وينطوي الشكل رقم (١) على مبالغة في التبسيط في أحد الجوانب المهمة : فهو يصور بث المعلومات عبر القنوات الرسمية ، ولكنه لا يعبر صراحة وبوضوح عن عمليات الاتصال غير الرسمي . ولا تثبت القنوات غير الرسمية ، بوجه عام ، معلومات تختلف عن تلك التي يتم بثها عبر القنوات الرسمية ؛ فكلاهما يثبت نتائج الخبرات أو البحوث نفسها إلا أن القنوات غير الرسمية تثبت المعلومات بشكل مختلف ، أو بالشكل نفسه ولكن بصورة أسرع أو في وقت مبكر كثيراً ، كما هو الحال مثلاً في توزيع المسودات أو الطباعات المبدئية Preprints . ولهذه القنوات أهميتها نظراً لأنها تثبت المعلومات بشكل أسرع من القنوات الرسمية ، على الأقل بالنسبة لمن يرتبطون ارتباطاً عضوياً بأحد الأوساط أو المجتمعات المهنية ، ولأنها أيضاً تثبت المعلومات إلى البعض الذين لا يفضلون القنوات الرسمية (*).

والهدف من هذه المناقشة الموجزة لدورة تداول المعلومات هو بيان دور مراكز المعلومات في الدورة الكاملة . وتتركز الاهتمامات الأساسية لهذا الكتاب في مجالات «الاقتناء في مراكز المعلومات» و «التنظيم والتعريف» و «التوزيع الثانوي» كما يصورها الشكل رقم (١) . ولما كانت بعض مراكز المعلومات تقوم أيضاً بإنتاج مرادف بيانات ثانوية ، يمكن أن تستخدم في تقديم مختلف الخدمات ، فإن الأنشطة الخاصة «بالنشر الثانوي» و «التوزيع الأولي» للمطبوعات الثانوية ، تدخل أيضاً ، جزئياً على الأقل ، في مجال هذا الكتاب . أما الأنشطة الأخرى التي تم تصويرها ، والتي تتصل بالتأليف ، والنشر الأولي والتوزيع الأولي ، واستيعاب المعلومات من جانب المستفيدين ، فلا تدخل في نطاق السيطرة المباشرة لمرافق المعلومات .

دور مرافق المعلومات :

الوظيفة الأساسية لأي مرفق من مرافق المعلومات هي العمل كهمزة وصل بين مجتمع معين من المستفيدين وعالم مصادر المعلومات في شكلها المطبوع وغير المطبوع (الشكل رقم ٢) .

وللاضطلاع بدوره كهمزة وصل يمارس مرفق المعلومات ثلاثة أنشطة أساسية (انظر الشكلين رقمي (١) و (٢)) ، وهي اقتناء الوثائق واختزانها ، وتنظيم هذه الوثائق والتعريف بها

(*) لمزيد من المعلومات حول أوجه الاختلاف بين القنوات الرسمية وغير الرسمية ، راجع : جاك ميدوز . آفاق الاتصال ومناخه في العلوم والتكنولوجيا ، ترجمة حشمت قاسم . القاهرة ، مكتبة غريب ، ١٩٧٩ م . (المترجم) .

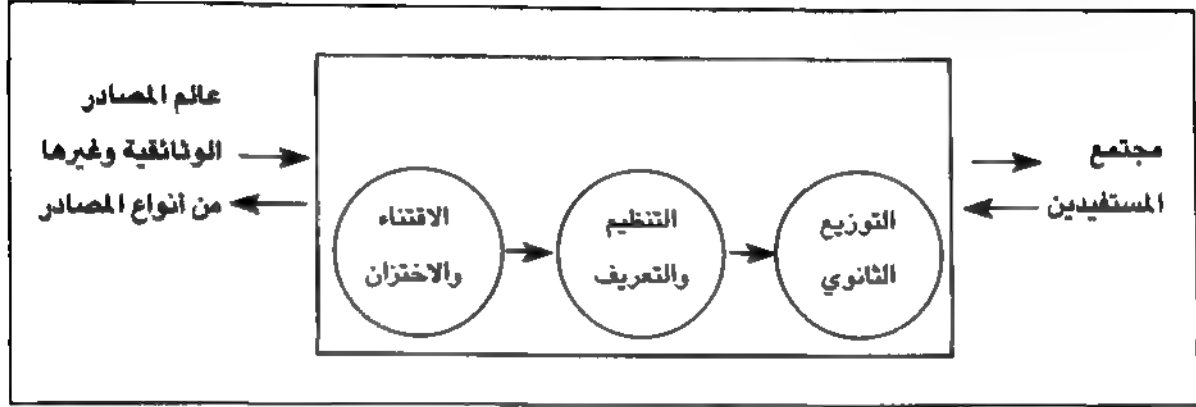
وتوزيع هذه الوثائق أو المعلومات المتصلة بها ، على المستفيدين ، عن طريق الإعارة ، وعمليات البحث في الإنتاج الفكري ، والاستنساخ الضوئي ، وغير ذلك من الخدمات .

وعادة ما يتم تحديد مجتمع المستفيدين على أساس جغرافي أو على أساس الانتماء المؤسسي أو على أساس الاهتمام الموضوعي ، أو بناءً على أكثر من أساس واحد من هذه الأسس ؛ فمجتمع المستفيدين بالنسبة للنظام الوطني للمعلومات ، هو مجتمع الباحثين العلميين وغيرهم من المهنيين العاملين في شتى أنحاء الدولة . وأهم أشكال مصادر المعلومات بالنسبة لمعظم مراكز المعلومات هو الشكل الوثائقي ، بأوسع معاني مصطلح « وثيقة » . ومهمة مرفق المعلومات هي جمع (وصل) كل من مجتمع المستفيدين ومصادر المعلومات معاً ، بشكل اقتصادي وفعال قدر الإمكان . وكدور سلبي إلى حد ما ، فإن مهمة مرفق المعلومات هي العمل على ضمان إتاحة أية وثيقة أو معلومة يحتاجها من ينتمون إلى مجتمع المستفيدين ، في متناول أيديهم ، بمجرد أن تدعو الحاجة إليها ، قدر الإمكان . وفي دوره الأكثر إيجابية يعمل مرفق المعلومات على تعريف مجتمع المستفيدين بالوثائق أو البيانات ، وربما عن طريق خدمات البحث في الإنتاج الفكري ، المصممة لإخطار هذا المجتمع ، وبشكل مستمر بما يصدر من إنتاج فكري جديد في مجالات اهتمامه (الإحاطة الجارية) (*) .

وينبغي لمرفق المعلومات الحديث الكفاء الفعال أن يكون قادراً على أن يجعل كل وثيقة في عالم الإنتاج الفكري المتوافر ، أو أية بيانات تشتمل عليها الوثائق ، في متناول أي عضو في مجتمع المستفيدين . ويعني ذلك أن يكون عالم المصادر الوثائقية في متناول مجتمع المستفيدين ، على درجات متفاوتة من السرعة واليسر . ونظراً لأنه لا يمكن لأي مركز للمعلومات أن يقتني كل شيء ، فإنه من المهم أن ينتقي أنسب الوثائق وأكثرها أهمية بالنسبة للمستفيدين من خدماته . ومن المهم أيضاً أن يكون مركز المعلومات قادراً على أن يقتني بسرعة أية وثيقة أخرى تدعو الحاجة إليها في مجتمع المستفيدين ، سواء عن طريق الشراء أو التصوير الضوئي أو تبادل الإعارة بين المكتبات . هذا بالإضافة إلى حاجة ما تضمه مقتنيات المركز من وثائق لأن تنظم تبعاً

(*) ترتبط الإيجابية والسلبية هنا بأوجه الاختلاف بين الأسس التي كانت تحكم أداء المكتبات التقليدية وتلك التي تحكم أداء مراكز التوثيق ثم مراكز المعلومات ؛ فقد كان من سياسة المكتبات انتظار المستفيد حتى يأتي التماساً للخدمة ، بينما كانت سياسة مراكز التوثيق ومرافق المعلومات الانتقال إلى المستفيد حيثما وجد ، عن طريق ما يعرف بالخدمات الدينامية . (المترجم) .

لمستويات الطلب المتوقع ، بحيث تكون تلك التي يمكن أن يشتد الطلب عليها أقرب ما تكون في متناول المستفيدين . وهكذا ينبغي أن تكون مجموعة المصادر الوثائقية في متناول المستفيدين من مركز المعلومات وفقاً لدرجات متفاوتة من سهولة المنال ، يمكن أن تكون بالنسبة لمؤسسات معينة على النحو التالي :



الشكل رقم (٢) : دور مرفق المعلومات كهمزة وصل .

- ١ - وثائق يكتنيها المركز ومتاحة على أرفف مفتوحة .
- ٢ - وثائق يكتنيها المركز ولكن في مخازن مغلقة .
- ٣ - وثائق في مستودعات بعيدة عن مقر المركز .
- ٤ - وثائق لا يكتنيها المركز .

وهذا التسلسل مبني على أساس تناقص سهولة المنال ، إلا أنه ينطوي على مبالغة في التبسيط ، نظراً لأنه من الممكن لوثيقة لا يكتنيها المركز أن تكون أيسر منالاً من وثيقة يكتنيها المركز فعلاً إلا أنها وضعت في غير مكانها ، أو حولت للتجليد أو أعيرت خارج المركز . أضف إلى ذلك أنه لا يمكن لجميع الوثائق التي لا يكتنيها المركز أن تتساوى في مدى سهولة المنال أو صعوبته ، نظراً لأنه من الممكن أن يكون للموقع الجغرافي للمركز أثره في سهولة الوصول إليها (وإن تضاعفت أهمية هذا العامل الجغرافي نتيجة لكثافة توافر أجهزة التصوير عن بعد وما يتصل بها من تقنيات) . وهناك بعض العوامل الأخرى التي تؤثر في سهولة المنال ، ومن بينها ما إذا كانت الوثيقة موجودة في أحد الفهارس الموحدة ، ومدى إمكان الاعتماد على خدمة الإيصال ومدى سرعة هذه الخدمة ، وهكذا .

فئات الحاجة إلى المعلومات :

تنقسم الاحتياجات والمطالب الأساسية للمستفيدين من مرافق المعلومات إلى فئتين عريضتين :

١ - الحاجة إلى العثور على وثيقة معينة نعرف اسم مؤلفها أو عنوانها ، والحصول على نسخة من هذه الوثيقة .

٢ - الحاجة إلى العثور على الوثائق التي تتناول موضوعاً بعينه ، أو يمكن أن تجيب عن سؤال معين .

ومن الممكن تسمية الفئة الأولى « بالحاجة إلى وثيقة معروفة » والفئة الثانية « بالحاجة الموضوعية » . وقدرة مركز المعلومات على تقديم الوثائق المعروفة هي « قدرته على الإمداد بالوثائق » ، أما قدرة المركز على استرجاع الوثائق المتخصصة في موضوع معين ، أو قدرته على الإجابة عن سؤال معين ، فهي « قدرته على استرجاع المعلومات » . وهاتان الوظيفتان ، الإمداد بالوثائق واسترجاع المعلومات هما أهم الأنشطة التي تقوم بها مرافق المعلومات . والعلاقة بين الوظيفتين وثيقة ، نظراً لأن كثيراً من طلبات الحصول على الوثائق المعروفة يمكن أن تكون ناشئة وبشكل مباشر عما سبقها من أنشطة استرجاع المعلومات .

هذا ، وتنقسم الاحتياجات الموضوعية ، بوضوح ، إلى ثلاثة أنواع :

١ - الحاجة إلى المعلومات المساعدة في حل مشكلة معينة أو لتيسير اتخاذ القرار .

٢ - الحاجة إلى المعلومات الأساسية حول موضوع معين .

٣ - الحاجة إلى المعلومات حول التطورات الجديدة في مجال معين .

وعادة ما يسمى النوع الأخير بالحاجة إلى الإحاطة الجارية ، إلا أنه ليس هناك مصطلح واحد يحظى بالقبول بوجه عام لوصف النوعين الآخرين . وعادة ما يتم ، في الممارسة الفعلية ، تلبية هذه الأنواع من الاحتياجات من خلال البحث أو التنقيب في رصيد الإنتاج الفكري السابق ، من جانب مرفق المعلومات ، استجابة لطلب معين من أحد المستفيدين ، فيما يسمى عادة « بالبحث الراجع » .

ويختلف البحث الراجع عن البحث للإحاطة الجارية في أكثر من وجه ؛ فالبحث الراجع أكثر تعمداً ، حيث يتعين أن تكون المبادرة من جانب المستفيد ، بينما يمكن في حالة الإحاطة الجارية أن تأتي المبادرة من مرفق المعلومات . هذا بالإضافة إلى أن البحث الراجع يميل لأن يكون أكثر تحديداً ، كما أنه يمكن أن يتطلب تغطية كم من الإنتاج الفكري أكبر كثيراً مما يتطلبه البحث للإحاطة الجارية ، حيث يرجع إلى الخلف عدة سنوات . أضف إلى ذلك أن نتائج البحث الراجع يمكن الحكم عليها من جانب المستفيد بمعايير أكثر تشدداً من تلك التي يحكم بها على نتائج الإحاطة الجارية .

ومن الممكن تقسيم أوجه الحاجة إلى البحث الراجع إلى عدة أنواع :

- ١ - الحاجة إلى عنصر بعينه من معطيات الحقائق : وهذه بعينها هي الأسئلة «المرجعية السريعة» التي تتناولها المكتبات . وعلى الرغم من أن الوثائق عادة ما تدخل في تلبية مثل هذه الحاجة ، فإنه ليس من الضروري أن يتلقى المستفيد الوثيقة ، حيث يمكن تقديم الإجابة عن السؤال بواسطة الهاتف .
- ٢ - الحاجة إلى الحصول على وثيقة واحدة أو أكثر حول موضوع معين ، ولكن بما لا يصل إلى إجمالي الإنتاج الفكري أو كل شيء متاح في أحد مراكز المعلومات : وهذا هو الحال عادة في المكتبات ، ومن أمثلته طلب عدد قليل من الوثائق الحديثة عن الأمطار الحمضية . ومن الفئات الخاصة هنا الحاجة التي يمكن تلبيتها كاملة بمجرد العثور على أول وثيقة من نوعية معينة ؛ فمن الممكن ، على سبيل المثال ، لفاحص براءات الاختراع أن يحتاج فقط إلى العثور على حالة واحدة لتطبيق سابق في الإنتاج الفكري لكي يرفض أو يفند الحق في إصدار براءة اختراع .
- ٣ - الحاجة إلى البحث الشامل ؛ البحث الذي يسترجع أكبر قدر ممكن من الإنتاج الفكري الصادر عن موضوع معين خلال فترة معينة : ويمكن أن يحتاج إلى مثل هذا البحث من يكون بصدد تأليف كتاب ، أو إعداد مراجعة علمية ، أو الباحث الذي يبدأ مشروع بحث . ومن الفئات الخاصة هنا البحث بهدف تأكيد خلو الإنتاج الفكري مما يتصل بموضوع معين ، أي عندما يريد المستفيد إثبات عدم نشر أي شيء عن الموضوع ، كما هو الحال مثلاً بالنسبة للمبتكر الذي يريد إثبات صلاحية ابتكار ما للتسجيل والحماية كاختراع .

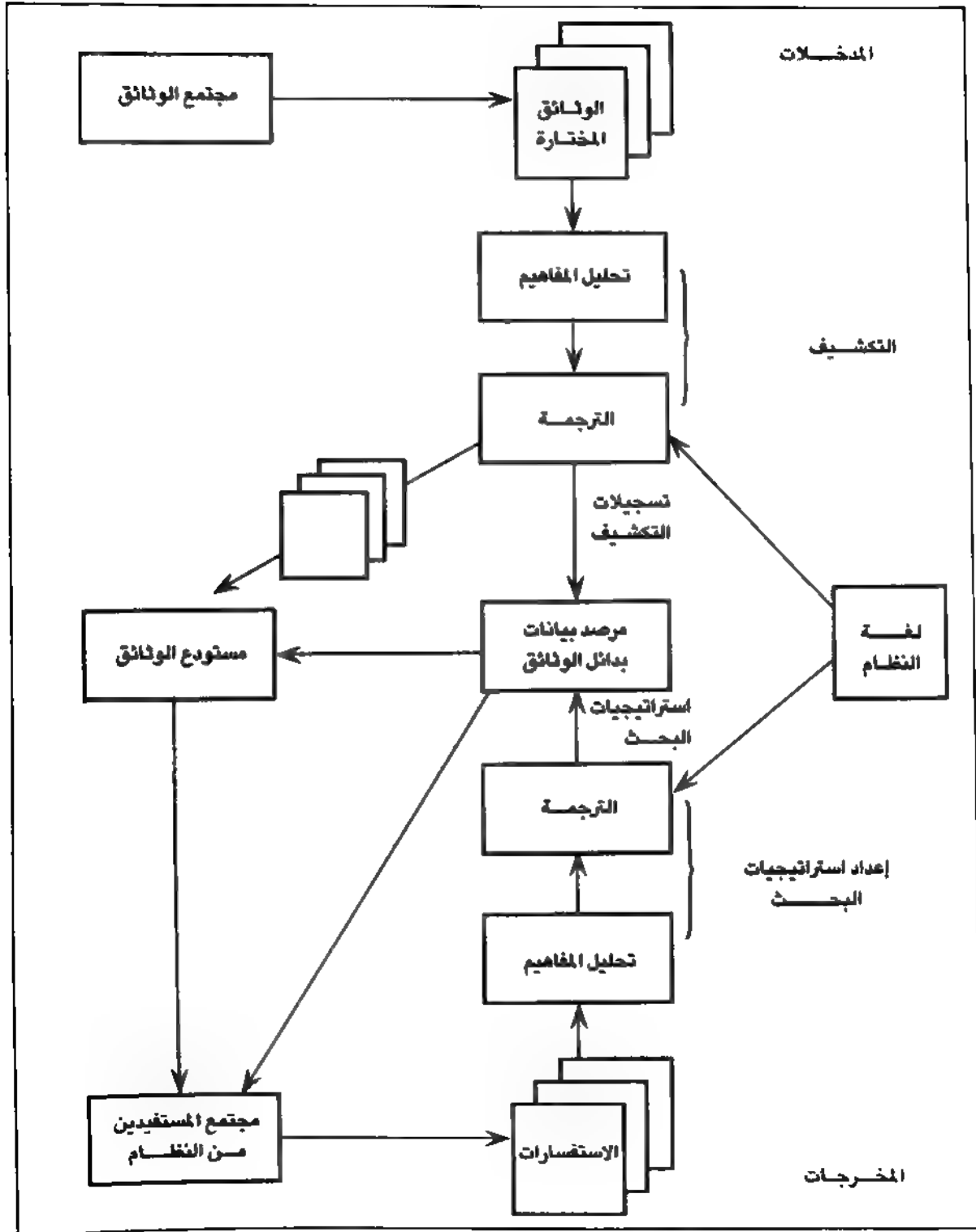
نظم استرجاع المعلومات :

يعبر الشكل رقم (٣) وبشكل مبسط عن أهم الأنشطة التي تمارسها أنواع كثيرة من مرافق المعلومات . فالمدخلات تتكون من الوثائق التي يقتنيها مرفق المعلومات . ويعني ذلك وجود معايير وسياسات للاختيار ، وتعني هذه المعايير والسياسات بدورها المعرفة التفصيلية الدقيقة باحتياجات المجتمع المزمع خدمته من المعلومات . والموضوع هو أبرز المعايير التي تحكم اختيار الوثائق ، إلا أن هناك معايير أخرى قد تكون لها أهميتها ، كنوعية الوثائق واللغة والمصدر . وفي مرافق المعلومات التي تتعامل أساساً مع مقالات الدوريات فإن معايير الاختيار عادة ما تركز على الدورية لا على المقالة . وهناك دوريات معينة يمكن أن تحظى بالتغطية وأخرى لا تتم تغطيتها (على الرغم من أن هناك بعض الدوريات التي يمكن أن تكشف جميع محتوياتها ، وأخرى تكشف انتقائياً) وعادة ما تكون التغطية في كثير من المرافق ، محكومة وإلى حد بعيد باعتبارات فعالية التكلفة ، وخاصة تلك المرافق التي تتعامل مع المجالات المغرقة في التخصص ، حيث لا تشمل التغطية سوى تلك الدوريات الأكثر اهتماماً بالنشر في الموضوع .

التحليل الموضوعي للوثائق :

تحتاج الوثائق بمجرد اقتنائها إلى أن تنظم ويتم التعريف بها ، حتى يكون من الممكن التحقق منها والوصول إلى مكانها استجابة لطلبات المستفيدين . وتشمل أنشطة التنظيم والتعريف كلاً من التصنيف والفهرسة والتكشيف الموضوعي والاستخلاص . وتنطوي عملية التكشيف الموضوعي على خطوتين فكريتين متميزتين : أولاهما تحليل المفاهيم أو تحليل محتوى الوثيقة ، والثانية ترجمة ناتج تحليل المفاهيم إلى لغة معينة (الشكل رقم ٣) . ونادراً ما يتم التمييز بين الخطوتين بوضوح . وهذا من سوء الحظ ، لأن كلاً من الخطوتين تفرض قيوداً مختلفة وتدخل عوامل مختلفة تؤثر في أداء النظام . ولضمان كفاءة تحليل المفاهيم يحتاج المكشف لأن يدرك ما تدور حوله الوثيقة ، كما ينبغي أن يكون على دراية باحتياجات المستفيدين . والتعرف على ما تدور حوله الوثيقة والأسباب المحتملة لاهتمام المستفيدين بها ، أي تحديد أهم عناصر محتوى الوثيقة هو جوهر التحليل الموضوعي . ومن الممكن تسجيل ناتج تحليل مفاهيم الوثيقة على الورق ، أو من خلال المنفذ الخاص بالنظام الآلي ، إلا أنه غالباً ما يكون في ذهن المكشف فقط .

وينبغي بعد ذلك ترجمة ناتج التحليل الموضوعي إلى لغة ما أو إلى لغة التكشيف . وينطوي ذلك في كثير من النظم على « لغة مقيدة » ، أي مجموعة محدودة من المصطلحات التي ينبغي أن تستخدم للتعبير عن المحتوى الموضوعي للوثائق . ويمكن لمثل هذه اللغة



الشكل رقم (٣) : الوظائف الأساسية لمعظم أنواع مرافق المعلومات .

أن تكون قائمة لرؤوس الموضوعات ، أو خطة للتصنيف ، أو مكنزاً ، أو مجرد قائمة بالكلمات المفتاحية أو العبارات « المقبولة » . أما اللغة غير المقيدة ، فعلى العكس لا تضع أية قيود على المصطلحات التي يمكن للمكشف أن يستخدمها . وعادة ما تعني اللغة غير المقيدة استخدام الكلمات أو العبارات التي ترد في الوثيقة التي يتم تكشيفها . ويشار إلى المصطلحات التي يستخدمها الكشف للتعبير عن المحتوى الموضوعي للوثائق ، سواء كانت من لغة مقيدة أو غير مقيدة ، يشار إليها عموماً في هذا الكتاب « بالمصطلحات الكشفية » .

وبمجرد انتهاء عملية التكشيف ، توضع الوثائق مرتبة في شكل من أشكال مستودعات الوثائق (*) ، بينما تدخل تسجيلات التكشيف في مستودع آخر وهو مرصد البيانات ، حيث تنظم بطريقة تيسر البحث فيها بالشكل المناسب استجابة لمختلف الاستفسارات . ويمكن لمرصد بيانات تسجيلات التكشيف ، أو تلك التسجيلات التي تشتمل على بيانات التحقق من الوثائق ، أو ما يسمى ببدائل الوثائق ، أن تكون في بساطة الملف البطاقي أو الكشاف المطبوع . إلا أنها من الأرجح في أيامنا هذه أن تكون في ملف قابل للقراءة بواسطة الآلات على شريط ممغنط أو على أسطوانة . ويمكن النظر إلى مرصد البيانات هذا باعتباره « كشاف » مستودع الوثائق .

وتؤدي التطورات التقنية الجارية إلى صعوبة التمييز بين مستودع الوثائق من جهة ومرصد بيانات بدائل الوثائق من جهة أخرى . فبدل الوثيقة الذي يتم البحث فيه قد لا يكون مجموعة من المصطلحات الكشفية ، وإنما عنصر نصي كالعنوان أو المستخلص وربما نص الوثيقة كاملاً . ويؤدي مرصد البيانات الذي يشتمل على نصوص الوثائق كاملة في شكل قابل للبحث ، وظائف كل من مستودع الوثائق ومرصد بيانات بدائل الوثائق أو تسجيلاتها .

استراتيجية البحث :

لا تختلف الخطوات التي ينطوي عليها جانب مخرجات النظام كثيراً عن تلك التي تنطوي عليها المدخلات . فمجتمع المستفيدين يتقدم بمختلف الاستفسارات إلى مرفق

(*) استخدم المؤلفان كلمة database لكل من الوثائق وبدائل الوثائق ، ولضمان وضوح الفارق بين العنصرين استعملنا مستودع الوثائق للدلالة على ملف الوثائق كاملة ، ومرصد البيانات للدلالة على ملف بدائل أو تسجيلات الوثائق . (المترجم) .

المعلومات ، حيث يقوم العاملون بإعداد استراتيجيات البحث اللازمة لهذه الاستفسارات . ومن الممكن النظر إلى إعداد استراتيجيات البحث أيضاً باعتباره ينطوي على كل من تحليل المحتوى أو تحليل المفاهيم والترجمة . فالخطوة الأولى تنطوي على تحليل الاستفسار لتحديد ما يبحث عنه المستفيد فعلاً ، بينما تنطوي الخطوة الثانية على ترجمة ناتج تحليل المفاهيم إلى لغة النظام . وتحليل مفاهيم أو محتوى الاستفسار ، ثم ترجمة ناتج هذا التحليل إلى لغة النظام هو « استراتيجية البحث » التي يمكن النظر إليها باعتبارها « بديل الاستفسار » تماماً كما يمكن النظر إلى تسجيلية الكشف باعتبارها بديل الوثيقة ، ووجه الاختلاف الوحيد بين الاثنين أن استراتيجية البحث عادة ما تشتمل على منطق واضح صريح ، أي تحديد مجموعة معينة من العلاقات المنطقية بين المصطلحات الكشفية ، في الوقت الذي قد لا يتم فيه بيان العلاقات المنطقية بين المصطلحات الكشفية ، في تسجيلية الوثيقة ، بصراحة ووضوح .

وبمجرد إعداد استراتيجية البحث ، فإنها « تضاهي » مقابل مرصد بيانات بدائل الوثائق . ويمكن لهذا النشاط أن ينطوي على بحث في الملفات البطاقية أو الكشافات المطبوعة ، أو المصغرات الفيلمية ، أو الأشرطة المغنطة أو الأسطوانات . ويتم استرجاع تسجيلات الوثائق التي تطابق استراتيجية البحث ، أي تلك التي تلبي الشروط المنطقية للبحث ، من مرصد البيانات وتقديمها للمستفيد . وتنتهي هذه العملية ، التي يمكن أن تكون تكرارية ، عندما يرضى المستفيد عن نتائج البحث ، ويمكن أن يعني ذلك في بعض الأحيان رضا المستفيد لأنه ما من شيء في مرصد البيانات يناسب احتياجاته على وجه التحديد .

وتوضح الخطوات المبينة في الشكل رقم (٣) موقف البحث المفوض ، أي الموقف الذي يفوض فيه من يحتاج إلى المعلومات مسئولية البحث في مرصد البيانات إلى أحد اختصاصيي المعلومات . أما في البحث غير المفوض فإن العملية عادة ما تتسم بالبساطة ، نظراً لأن المستفيد يتجه مباشرة إلى مرصد البيانات . إلا أنه حتى في هذا الموقف ، يتعين على المستفيدين تحليل مفاهيم احتياجاتهم من المعلومات ، وترجمة نتائج تحليلاتهم إلى لغة النظام . وعند البحث في كثير من أنواع النظم ، لا يتم بناء استراتيجيات البحث بمنأى عن مراصد البيانات أو بمعزل عن عملية البحث نفسها . فمن الممكن للاستراتيجية الخاصة بالبحث في الفهرس البطاقي أو في الكشاف المطبوع أو في مرصد

البيانات الإلكترونية (في شكل الأسطوانات الضوئية المكتنزة ، أو المتاح على الخط المباشر عبر شبكات الاتصالات بعيدة المدى) من الممكن أن تتطور تفاعلياً وإيعازياً ، أي أن أنشطة تحليل المفاهيم والترجمة تتزامن تقريباً مع أنشطة البحث في الملفات . ومع ذلك ، فإنه من الأوفق فصل تحليل المفاهيم عن الترجمة ، حتى في هذا الموقف .

ووجه الاختلاف الحقيقي الوحيد بين خدمة البحث الراجع وخدمة الإحاطة الجارية كالبث الانتقائي للمعلومات مثلاً ، أن استراتيجيات البحث ، أو « سمات اهتمامات المستخدمين » في الأخيرة تمثل اهتمامات البحوث الجارية أو الآنية للمستخدمين من النظام . وهذه الاستراتيجيات أو السمات تتم مضاهاتها مقابل بدائل أو تسجيلات الوثائق الواردة على أساس منتظم ، أي كلما تم تجديد مرصد البيانات ، تقدم نتائج عمليات المضاهاة هذه للمستخدمين على فترات منتظمة .

استرجاع المعلومات في مقابل ما يتصل به من أنشطة :

هناك بعض الاستفسارات الخاصة بوثائق معينة يعرف المستخدمون أسماء مؤلفيها أو عناوينها ، ولا تتعلق بالمعلومات حول موضوع معين . وتقدم هذه الاستفسارات (الشكل رقم ٣) إلى مستودع الوثائق مباشرة ، ووفقاً لمدخل المؤلف أو مدخل العنوان في الكشافات أو فهرس المقتنيات ، أو وفقاً لمدخل آخر كرقم التقرير أو رقم براءة الاختراع مثلاً . وقدرة مركز المعلومات على تقديم الوثائق المطلوبة هي « قدرته على الإمداد بالوثائق » (*).

ولم نحاول حتى الآن تقديم تعريف محدد « لنظام استرجاع المعلومات » على الرغم من تعرضنا لمجال هذا المصطلح . « فاسترجاع المعلومات » ، كما يستعمل في الغالب الأعم من الحالات مرادف « للبحث في الإنتاج الفكري » ، فهو عملية البحث في مجموعة من الوثائق (وفقاً لأوسع معاني المصطلح « وثيقة ») للتحقق من تلك الوثائق التي تتناول موضوعاً بعينه . ومن ثم فإن أي نظام مصمم لتيسير عملية البحث في الإنتاج الفكري هذه يمكن أن يسمى ، وبلا تجاوز ، نظاماً لاسترجاع المعلومات . والفهرس الموضوعي بالمكتبة أحد أنواع هذه النظم ، وكذلك أيضاً الكشاف الموضوعي المطبوع .

(*) فضلاً عما تنطوي عليه هذه العبارة من تكرار ، فإن استعمال المؤلفين لعبارة « تقدم هذه الاستفسارات إلى مستودع الوثائق مباشرة » ليس له ما يبرره ، لأن الرجوع المباشر إلى مستودع الوثائق لا يتم إلا في حالة معرفة أماكن الوثائق المطلوبة في هذا المستودع . (المترجم) .

وعادة ما تتكون مخرجات نظام استرجاع المعلومات من إشارة ورقية واحدة أو أكثر، مصحوبة في بعض الأحيان بمعلومات إضافية كالمستخلصات أو قوائم المصطلحات التي كُشفت تحتها الوثائق . وتقدم بدائل الوثائق هذه إلى من طلب إجراء البحث . ويمكن لذلك الشخص أن يطلب من مركز المعلومات أو غيره من المرافق تزويده بالوثائق المشار إليها . ويحدث في بعض الأحيان أن يتجنب مركز المعلومات الخطوة الوسيطة ، ويقدم مباشرة للمستفيد الوثيقة الفعلية أو عينة من الوثائق التي يراها اختصاصيو المعلومات الأقرب إلى تلبية حاجة المستفيد . ويحدث في بعض الأحيان الجمع بين كل من وظيفة استرجاع المعلومات ووظيفة تقديم الوثائق في نظام واحد . فمن الممكن على سبيل المثال للنظام الآلي أن يشتمل على النصوص الكاملة للوثائق ، بحيث تكون مخرجات البحث ناتجاً مطبوعاً للوثائق لا لبدائلها أو تسجيلاتها . إلا أن معظم نظم استرجاع المعلومات تقدم تسجيلات الوثائق ، أما تقديم الوثائق فنشاط مستقل تماماً ، يمكن للمؤسسات التي تقوم بإجراء عمليات البحث في الإنتاج الفكري أن تنهض به أو لا تنهض .

ومن الواضح أن « استرجاع المعلومات » ليس بالمصطلح المرضي للدلالة بوجه خاص على نوعية النشاط الذي يستعمل معه عادة ؛ فنظام استرجاع المعلومات لا يسترجع معلومات ، والمعلومات ، في الواقع ليست بالشيء الذي يمكن لمسه ، ولا يمكن رؤيتها أو سماعها أو الإحساس بها . فنحن نصبح « على علم Informed » إذا ما طرأ تغير ما على حالتنا المعرفية في موضوع ما . وإعطاء المستفيد وثيقة أو إشارة ورقية لوثيقة عن الليزر لا يجعله على علم في موضوع الليزر . ولا يمكن لانتقال المعلومات أن يتم فعلاً إلا إذا اطلع المستفيد على الوثيقة وفهمها . فالمعلومات إذن هي الشيء الذي يغير الحالة المعرفية للشخص في موضوع ما . وقد لا يكون هذا تعريفاً دقيقاً جداً ، إلا أنه أفضل ما نستطيع تقديمه ، وهو مناسب على الأقل لأغراضنا .

ومما يزيد في تأكيد عدم صلاحية المصطلح « استرجاع المعلومات » إعادة النظر في الشكل رقم (١) . فأنشطة نظم استرجاع المعلومات تنتهي بالدائرة الموسومة « بالتوزيع الثانوي » ، حيث المواد التي يتم توزيعها هي الوثائق أو تسجيلاتها . أما تداول المعلومات ، إذا قدر له أن يحدث على الإطلاق ، فلا يحدث إلا في الدائرة الموسومة « بالاستيعاب من جانب المستفيدين » . ولا تدخل مرحلة الاستيعاب ، بدورة تداول المعلومات ، في نطاق سيطرة مراكز المعلومات .

وهناك نشاط آخر من أنشطة مراكز المعلومات ، يسمى أحياناً « الرد على الاستفسارات » ، يمكن النظر إليه باعتباره أحد أشكال استرجاع المعلومات . وتحاول خدمة الرد على الاستفسارات تقديم الإجابة المباشرة لأسئلة معينة ، مثل : ما ارتفاع كذا...؟ ما درجة انصهار كذا...؟ ما عنوان كذا...؟ لا مجرد الإشارة ببساطة إلى الوثائق التي يمكن أن تقدم الإجابة . وتقدم كثير من المكتبات وغيرها من ضروب مرافق المعلومات خدمة الرد على الاستفسارات ، التي تسمى أحياناً بخدمة المراجع السريعة أو الإرشاد بالحقائق . ويمكن للرد على الاستفسارات أن يكون المرحلة الثانية في نشاط أوسع لاسترجاع المعلومات ، حيث تنطوي الخطوة الأولى على استخدام أحد أنواع نظم استرجاع المعلومات كالفهرس الموضوعي للمكتبة ، أو أحد الكشافات المطبوعة ، أو ربما كشاف نهاية الكتاب ، وذلك للتحقق من الوثائق التي يمكن أن تقدم الإجابة عن سؤال معين . أما الخطوة الثانية فتنتطوي على اقتطاف الإجابة من الوثيقة وتقديمها للمستفيد .

ولقد تم تطوير العديد من النظم الإلكترونية للرد على الاستفسارات . وتتلقى هذه النظم الأسئلة باللغة الطبيعية ، وإن كان ذلك يمكن أن يتطلب بناءً نظامياً موحداً ، ثم تقدم الإجابة مطبوعة أو على الشاشة . ونظراً لتعدد تصميم مثل هذه النظم ، فإن ما تم تطويره منها حتى الآن يقتصر على مجالات معرفية محدودة كنتائج مباريات البيزبول ، أو قواعد الطرق السريعة للولايات . ومن الممكن تسمية النظام الذي يخزن البيانات الفيزيائية أو الكيميائية أو غيرها من ضروب البيانات ، ويجيب عن الأسئلة اعتماداً على البيانات المخزنة ، « بنظام استرجاع البيانات » . ويمكن تسمية البيانات أو المعطيات المخزنة « ببنك البيانات » . ومن الأمثلة على ذلك بيانات التعداد ، أو بيانات الخواص الحرارية الفيزيائية ، أو البيانات الخاصة بطاقات الوضع الداخلية للذرات *Interatomic Potentials* .

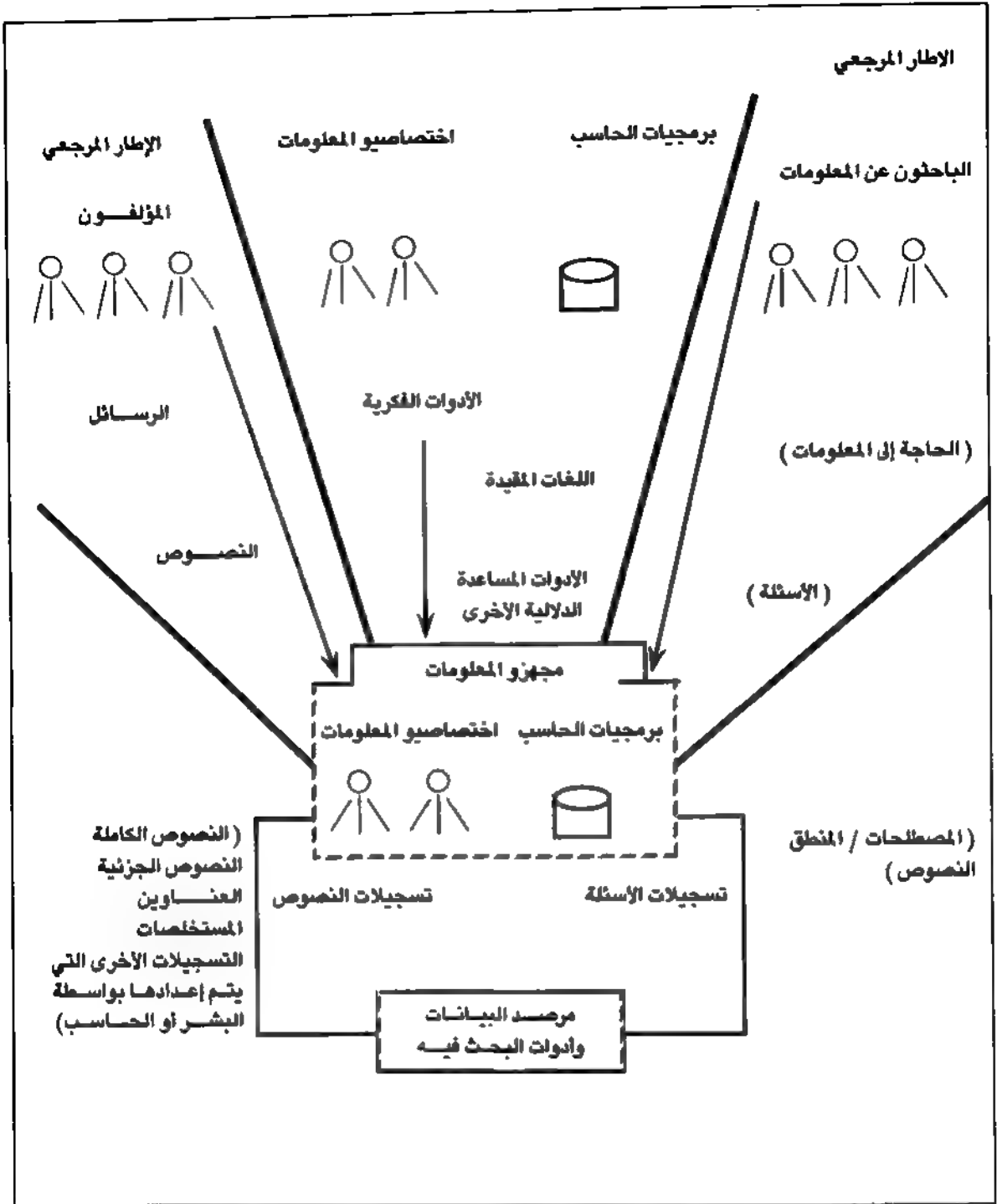
وهناك ضرب آخر من نظم استرجاع المعلومات ، هي النظم التي تخزن النص الكامل لمجموعة من الوثائق ، والتي يمكن أن تسترجع تلك القطاعات من النص التي تطابق استراتيجيات البحث التي تعبر عن احتياجات المستفيدين . فمن الممكن للنظام الإلكتروني أن يخزن مجموعة من النصوص القانونية كدستور إحدى الولايات ، ويسمح باسترجاع تلك الفقرات من النص ، والتي ترد بها تجميعات معينة من الكلمات ، كالكلمة التي تدل مثلاً على الطفل والكلمة التي تدل على القوة البدنية .

ونظم الرد على الاستفسارات ، ونظم استرجاع البيانات أو المعطيات ، ونظم استرجاع النصوص ، كلها أشكال معتمدة لنظم استرجاع المعلومات . وعلى الرغم من الإشارة إلى نظم الرد على الاستفسارات ، ونظم استرجاع المعطيات ، أحياناً في هذا الكتاب ، فإن التركيز الأساس ينصب على النظم التي تسترجع تسجيلات الوثائق أو نصوص الوثائق استجابة للأسئلة الموضوعية (*) .

مشكلات استرجاع المعلومات :

لقد قدم الشكل رقم (٣) صورة مبسطة إلى حد ما لمشكلة استرجاع المعلومات ، ونقدم في الشكل رقم (٤) تصويراً أكثر تفصيلاً للمشكلة . والمشكلة في جوهرها هي مضاهاة الحاجة إلى المعلومات مقابل الرسائل ، ولا يمكن لذلك أن يتحقق إلا بشكل غير مباشر ؛ فمعظم الرسائل (ما يريد المؤلفون إيصاله) تظهر في شكل نصوص (بعضها في شكل تصويري ، أو سمعي ، أو في أي شكل آخر غير نصي) بينما الحاجة إلى المعلومات يتم التعبير عنها في شكل استفسارات تقدم إلى نوع ما من مرافق المعلومات . وينشئ مرفق المعلومات تسجيلات للنصوص (يمكن أن تتكون من النصوص الكاملة ، كما هو الحال مثلاً في الاستنساخ الإلكتروني المدقق للكلمات المطبوعة على الورق ، أو تتكون من أجزاء من النصوص ، أو من أي شكل آخر يمكن أن يتم إنشاؤه بواسطة البشر أو بواسطة الآلات) ثم يختزن هذه التسجيلات في مرصد البيانات ، ويوفر بعض الوسائل التي تكفل إمكانية البحث في هذه التسجيلات . ومن الممكن اختزان مرصد البيانات في شكل ورقي أو في شكل صور مصغرة ، أو في شكل إلكتروني ، أما « الوسيلة » التي تكفل إمكانية البحث فيه فيمكن أن تكون في بساطة الفهرس البطاقي أو الكشاف المطبوع ، كما يمكن أن تكون متطورة في شكل مجموعة من برمجيات الحاسب الآلي . كما يقوم مرفق المعلومات أيضاً بإعداد تسجيلات للأسئلة أو الاستفسارات (صيغ البحث أياً كان نوعها ، التي تتكون من المصطلحات ، أو المصطلحات مع بيان ما بينها من علاقات منطقية ، أو من عبارات نصية ، أو من « وثائق » ، حيث يمكن على سبيل المثال ، لنظام ما أن يكفل للباحث إمكانية إدخال بيانات وثيقة ، يعرف فعلاً أنها مناسبة ، ثم يبحث عن

(*) تسمى هذه الفئة الأخيرة عادة بنظم استرجاع الوثائق ، بينما تسمى نظم الرد على الاستفسارات ونظم استرجاع المعطيات بنظم استرجاع الحقائق . (المترجم) .



الشكل رقم (٤) : المشكلات الأساسية لاسترجاع المعلومات .

الوثائق الأخرى التي تشبهها (*) ومضاهاة هذه التسجيلات مقابل مرصد البيانات ، واسترجاع تلك التسجيلات الخاصة بالنصوص التي تتطابق تماماً مع تسجيلات الاستفسارات .

(*) كما هو الحال في الاسترجاع في مرصد بيانات كشافات الاستشهاد المرجعي . (المترجم) .

وقد لا تكون النصوص تسجيلات أو بدائل مثالية أو كاملة للرسائل (وعلى الرغم من أن هذه تدخل في صميم مشكلات الاتصال ، فإنها لا ينظر إليها عادة كأحدى مشكلات استرجاع المعلومات) كما أن بدائل النصوص (كتلك التي يعدها المكشفون أو المستخلصون مثلاً) قد تكون أيضاً غير كاملة أو غير مثالية . وبناء على المبدأ نفسه ، فإن الأسئلة أو الاستفسارات قلما تكون بدائل أو تسجيلات مثالية للحاجة إلى المعلومات ، كما أن صيغ أو عبارات البحث قد لا تكون أيضاً بديلاً أو تعبيراً مثالياً عن الاستفسارات . أضف إلى ذلك أن الإطار المرجعي للمستفيد قد لا يتفق تمام الاتفاق مع الإطار المرجعي لاختصاصيي المعلومات ، أو الإطار المرجعي للمؤلفين أنفسهم . وعلى ذلك ، فإنه من الممكن النظر إلى مشكلة استرجاع المعلومات باعتبارها في جوهرها مشكلة محاولة مضاهاة التقديرات التقريبية Approximations للحاجة إلى المعلومات بالتقديرات التقريبية للرسائل . ولا عجب إذن أن لا تأتي النتائج مرضية تماماً على الدوام .

وترى بيتس (Bates 1986) أن مشكلة استرجاع المعلومات أكثر تعقداً مما تبدو لأول وهلة ، وتصفها بأنها « غامضة غير محددة المعالم » و « ترجيحية أو احتمالية » . ويبدو أن الاتجاه السائد الآن هو التركيز على جانب المخرجات في الأنشطة (التعبير عن الحاجة إلى المعلومات والاستفسارات) أكثر من التركيز على جانب المدخلات (التعبير عن الرسائل والنصوص) والاعتقاد الكامن وراء ذلك هو أن جانب المخرجات أكثر « تعقداً » . ويشير بلكن (Belkin 1980) وبلكن وآخرون (Belkin et al. 1982 A,B) في الواقع إلى مضاهاة « الحالة المعرفية غير السوية » للسانل أو من يقدم الاستفسار ، « بالحالة المعرفية الأكثر تماسكاً أو المترابطة منطقياً » للمؤلفين . إلا أن دور المكشف ، أي التنبؤ بأنواع الاستفسارات التي يمكن لوثيقة معينة أن تكون إجابة مفيدة لها ، ليس من الضروري أن يكون أبسط من دور وسيط البحث ، وهو إدراك نوعيات الوثائق التي يمكن أن تلبي حاجة أحد المستفيدين في وقت معين .

وإذا كان الأمر كذلك فعلاً ، فإن الشكل رقم (٤) يبين أيضاً إمكان استخدام العمليات الخوارزمية في مختلف أنشطة استرجاع المعلومات ، كبديل للتجهيز الفكري البشري ؛ فمن الممكن استخدام الحاسبات الآلية في الكشف الآلي والاستخلاص الآلي ، وكذلك في العمليات الأخرى التي تنطوي على تشكيل فئات من الوثائق والمصطلحات ، وفي صياغة استراتيجيات البحث ، وفي تكوين شبكات لما بين المصطلحات من مصاحبات . وكما يفهم

من الشكل ، فإنه يمكن للحاسبات الآلية مساعدة البشر ، أو أن تقوم مقامهم إلى حد ما في جميع العمليات تقريباً التي تشكل نشاط استرجاع المعلومات بكل مكوناته .

مكونات نظم استرجاع المعلومات :

تتضح المكونات الأساسية لنظام استرجاع المعلومات في الشكل رقم (٣) ، حيث يتكون النظام من ستة نظم فرعية أساسية :

- ١ - النظام الفرعي لاختيار الوثائق .
- ٢ - النظام الفرعي للتكشيف .
- ٣ - النظام الفرعي للغة .
- ٤ - النظام الفرعي للبحث .
- ٥ - النظام الفرعي للتفاعل بين المستفيد والنظام (مواجهة المستفيد مع النظام) .
- ٦ - النظام الفرعي للمضاهاة ، أي النظام الفرعي الذي يقوم فعلاً بمضاهاة تسجيلات الوثائق مقابل تسجيلات الاستفسارات .

ويعد النظام الفرعي للمضاهاة ، في معظم أوساط العمل الفعلية ، أقل النظم الفرعية الستة أهمية ، لأنه ليس له تأثير مباشر على فعالية النظام ككل ، أي على ما إذا كان من الممكن للنظام أن يسترجع الوثائق التي تلبي حاجة المستفيدين إلى المعلومات (على الرغم مما لكفاءة النظام الفرعي للمضاهاة من أثر كبير واضح على اقتصاديات النظام وكفاءته بوجه عام إذا ما قيست على أساس زمن الاستجابة) . وفي النظم الإلكترونية التقليدية لا يسهم الحاسب وبشكل مباشر إلا في عملية المضاهاة ، حيث يعمل كوسيلة مضاهاة عملاقة . وقلما يسهم الحاسب الآلي ، في كثير من النظم ، وقد لا يسهم مباشرة على وجه الإطلاق في اختيار الوثائق ، أو تكشيف الوثائق ، أو التحكم في اللغة المستخدمة في التكشيف والبحث ، أو في إعداد استراتيجيات البحث ، أو في التفاعل مع المستفيدين من النظام لأغراض التفاوض حول الاستفسار مثلاً . وكل هذه أنشطة فكرية ، يقوم بها البشر في معظم النظم القائمة ، وتحكم فعالية النظام . وهناك العديد من النظم التجريبية التي يتم فيها استخدام الآلات بكامل طاقتها في الأنشطة المختلفة ونتناول هذه الأنشطة في الفصل الحادي عشر .

ومن الممكن تقسيم أهم العوامل المؤثرة في فعالية نظم استرجاع المعلومات إلى فئتين: العوامل الخاصة بمرصد البيانات ، والعوامل المرتبطة بالإفادة من مرصد البيانات. أما العوامل الجوهرية الخاصة بمرصد البيانات ، والتي يمكن النظر إليها أيضاً باعتبارها « عوامل المدخلات » فهي :

١ - الوثائق التي يغطيها النظام .

٢ - مدى الدقة والاكتمال في التعرف على المحتوى الموضوعي للوثائق والتعبير عنه في عملية الكشف .

٣ - مدى صلاحية لغة النظام للتعبير عن المحتوى الموضوعي .

وهناك أيضاً ثلاثة عوامل ترتبط بالإفادة من النظام أو بالخرجات :

١ - مدى قدرة العاملين بمركز المعلومات على إدراك حاجة المستفيدين إلى المعلومات (تفاعل المستفيد مع النظام) .

٢ - مدى قدرة العاملين بالمركز على تحويل احتياجات المستفيدين إلى استراتيجيات بحث .

٣ - مدى صلاحية لغة النظام للتعبير عن الاهتمامات الموضوعية للمستفيدين . والمكونات الفكرية لنظم استرجاع المعلومات ، وما لهذه المكونات من أثر في أداء النظام ، هي نقاط التركيز الأساسية لهذا الكتاب .

المكتبة كنظام لاسترجاع المعلومات :

مرفق استرجاع المعلومات هو ذلك المرفق الذي يستجيب لطلب المستفيد للمعلومات حول موضوع معين ، بالبحث في مرصود البيانات المطبوعة أو الإلكترونية ، للتحقق من الوثائق التي يمكن أن تتناول هذا الموضوع .

ولم يصبح هذا الضرب من خدمات المعلومات ، في متناول معظم أنواع المكتبات على نطاق واسع نسبياً ، إلا في غضون العشرين عاماً الماضية . فقد كانت هذه الخدمات لا تتوافر من قبل إلا في مكتبات متخصصة معينة ، وخاصة تلك المكتبات المتخصصة في المجالات الصناعية . أما المكتبات العامة والمدرسية والأكاديمية فقد كانت تعاني شح

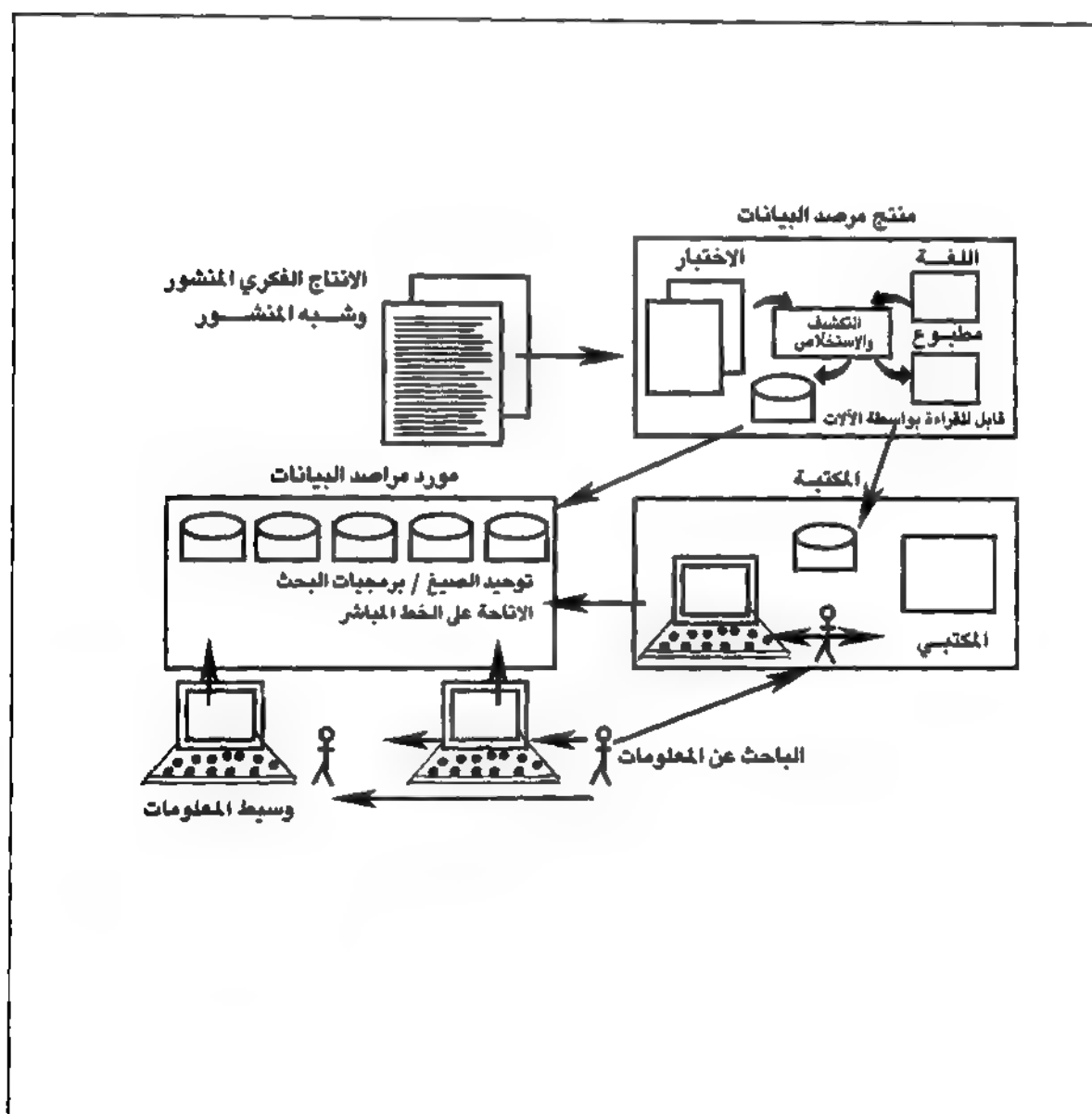
الموارد الذي لم يكن يشجعها إلا على إجراء أبسط عمليات البحث الوراقى . ومن ثم فإنها كانت توجه المستفيدين نحو المصادر المطبوعة المناسبة ، حيث كان بإمكانهم إجراء عمليات البحث الخاصة بهم بأنفسهم ، كما كانت ترشدهم إلى كيفية الإفادة من هذه المصادر إذا دعت الضرورة .

ولقد تغير هذا الموقف تغيراً جوهرياً منذ بداية سبعينيات القرن العشرين ؛ فقد أصبح استخدام شبكات الخط المباشر للبحث في مرصد البيانات الوراقية ، الآن مألوفاً في المكتبات الأكاديمية والمتخصصة على اختلاف أحجامها ، فضلاً عن توافره في بعض المكتبات العامة الكبرى . كذلك أصبحت مرصد البيانات التي توزع في شكل الأسطوانات الضوئية المكتنزة CD-ROM من الأمور المألوفة في جميع أنواع المكتبات ، كما تقوم بعض الجامعات الآن بتحميل مرصد البيانات الوراقية الضخمة على حاسباتها الآلية ، حتى يمكن التعامل معها عبر المنافذ المنتشرة في شتى أنحاء الحرم الجامعي .

وتقوم الآن بين الأفراد والمؤسسات العاملة في مجال البحث على الخط المباشر مجموعة معقدة من العلاقات (الشكل رقم (٥)) . ويضطلع منتج مرصد البيانات بالأدوار الأساسية في العملية برمتها كجامع وناشر لمرصد البيانات . وينطوي التجميع على اقتناء الوثائق المنشورة في نطاق المجال المحدد لمرصد البيانات (ويعني ذلك تطبيق معايير اختيار دقيقة) وتجهيز هذه الوثائق لإعداد بدائل وراقية (تسجيلات) في إطار مرصد البيانات . ويمكن لهذا التجهيز أن يشمل الفهرسة الوصفية والتكشيف الموضوعي (وربما باستخدام المصطلحات المستقاة من إحدى اللغات المقيدة كالمكنز مثلاً) ، كما يشمل إعداد المستخلصات في بعض الأحيان . إلا أنه يمكن في بعض الأحيان الاقتصاد إلى أقصى حد ممكن في العمليات الفكرية للتجهيز ، حيث تستخدم الكلمات المفتاحية الواردة في العناوين والمستخلصات كمدخل بدلاً من المصطلحات الكشفية التي يتم تعيينها بواسطة المكشفين . وغالباً ما توزع الآن مرصد البيانات في شكلين ؛ شكل قابل للقراءة بواسطة الآلات (إلكتروني) وشكل مطبوع (بمستخلصات أو بدون مستخلصات) مناظر تقريباً للشكل الإلكتروني .

ويقوم موردو مرصد البيانات بتجميع مرصد البيانات القابلة للقراءة بواسطة الآلات ، حيث طوروا البرمجيات اللازمة لتحويل جميع مرصد البيانات إلى صيغة تجهيز

موحدة ، وإتاحتها على الخط المباشر عن طريق مختلف شبكات الاتصالات بعيدة المدى ، وجعلها قابلة للاستجواب من جانب المستخدمين عن بعد . وعادة ما تتعامل المكتبات مع مراصد البيانات هذه عن طريق واحد أو أكثر من متعهدي مراصد البيانات، وإن كان من الممكن في حالات قليلة لمنتجي مراصد البيانات إتاحة إمكانية التعامل أيضاً مع مراصدهم على الخط المباشر عن طريق أجهزة الحاسبات الآلية الخاصة بهم . أضف إلى ذلك تزايد اتجاه المكتبات بوجه عام نحو الحصول على مراصد البيانات في شكل إلكتروني (الأسطوانات الضوئية المكتنزة) أو حرص المؤسسات الراعية للمكتبات على الحصول على مراصد البيانات وإتاحة فرصة تعامل المستخدمين معها على الخط المباشر.



الشكل رقم (٥) : موقف البحث على الخط المباشر في الولايات المتحدة الأمريكية .

ويمكن لمن يلتمس المعلومات أن يطلب من المكتبي إجراء البحث . وعادة ما يتم ذلك على الخط المباشر ، وإن كان من الممكن أن ينطوي أيضاً على التعامل مع مراصد البيانات المطبوعة أو المسجلة على الأسطوانات الضوئية المكتنزة . كذلك يمكن للمستفيدين التوجه إلى المكتبة لإجراء عمليات البحث الخاصة بهم في الأدوات المطبوعة ، أو استخدام منافذ المكتبة للتعامل مباشرة مع مراصد البيانات . ومع تزايد توافر المنافذ بالمكاتب والمنازل ، أصبح بإمكان المزيد من المستفيدين إجراء عمليات البحث الخاصة بهم على الخط المباشر ، دون مساعدة من المكتبيين . وتفضل بعض المكتبات المتخصصة الآن فعلاً تدريب المستفيدين على إجراء عمليات البحث الخاصة بهم بدلاً من إجراء هذه العمليات نيابة عنهم . وبدلاً من التوجه إلى المكتبة أو إجراء عمليات البحث بأنفسهم يمكن لبعض الأفراد والمؤسسات أن يفضلوا الاستفادة من خدمات وسطاء المعلومات الذين يقومون بإجراء عمليات البحث في مراصد البيانات مقابل رسوم يتقاضونها .

ومن الممكن بوجه عام النظر إلى المكتبة باعتبارها نظاماً لاسترجاع المعلومات . وبعبارة أخرى فإن الأنشطة التي يصورها الشكل رقم (٣) هي تلك الأنشطة التي تمارسها معظم المكتبات ؛ فهي على سبيل المثال تقتني المطبوعات على اختلاف أنواعها لتنمية مجموعاتها ، أو مستودع الوثائق ، كما أنها تعد مرصداً للبيانات من تسجيلات الوثائق ، وهو هنا فهرس المكتبة . ويبحث المستفيدون من المكتبة في هذا الفهرس للتحقق من الوثائق التي تتناول موضوعاً بعينه ، كما يمكن للمكتبيين إجراء عمليات البحث هذه نيابة عنهم . وواقع الأمر أن أنشطة المكتبات مماثلة تماماً لأنشطة ناشري مراصد البيانات . وعلى ذلك ، فإن الشكل رقم (٣) يمكن أن ينطبق على المكتبات ، كما ينطبق مثلاً على مؤسسة المستخلصات الكيميائية Chemical Abstracts Service ، أو على بعض مراكز المعلومات المتخصصة في الأجهزة الحكومية .

ولأسف فإن الفهارس البطاقية (أو الفهارس المطبوعة بالنسبة لهذا الأمر) يكتنفها مظاهر قصور لا مراء فيها كأدوات للاسترجاع الموضوعي ، إلا أنه منذ أصبحت الفهارس المتاحة على الخط المباشر مألوفة في المكتبات ، انبعث الاهتمام بالمدخل الموضوعي بوضوح في جميع أوساط مهنة المكتبات ، نظراً لأن هذه الفهارس المتاحة على الخط المباشر تتيح إمكانية إجراء عمليات البحث بشكل أكثر تطوراً ، كما تكفل مداخل موضوعية أكثر فعالية بوجه عام . ولا تختلف العوامل المؤثرة في الاسترجاع الموضوعي

عن طريق فهرس المكتبات عن تلك التي تؤثر في الاسترجاع الموضوعي عن طريق أي نوع آخر من مراصد البيانات الوراقية على الإطلاق .

ومما يدعو للأسف أن المصطلحات التي تستخدم عادة للدلالة على مختلف جوانب المعالجة الموضوعية في إطار مهنة المكتبات ، لا تتفق تماماً والمصطلحات التي يستخدمها الآخرون ، كناشري مراصد البيانات الوراقية مثلاً . وأحياناً ما نجد في الإنتاج الفكري للمكتبات وعلم المعلومات تمييزاً بين مصطلحات « الفهرسة الموضوعية » و « التكشيف الموضوعي » و « التصنيف » الثلاثة .

ويدل مصطلح الفهرسة الموضوعية عادة على استعمال رموز الموضوعات للتعبير عن المحتوى الإجمالي للوحدات الوراقية الكاملة (كالكتب والتقارير والدوريات إلى آخر ذلك من أنواع أوعية المعلومات) في فهرس المكتبة .

أما مصطلح التكشيف الموضوعي فيستخدم بأكثر من معنى ؛ فمن الممكن أن يستخدم للدلالة على التعبير عن المحتوى الموضوعي لأجزاء من الوحدات الوراقية الكاملة ، كما هو الحال في كشف نهاية الكتاب أو كشف المقالات التي تصدر في الدوريات . وهكذا يمكن للمكتبة أن تفهرس الكتاب تحت رأس موضوع « المكتبات » للدلالة على إجمالي المحتوى الموضوعي ، ولا يتم الكشف عن دقائق محتويات الكتاب إلا في الكشف الموضوعي . وهذا التمييز بين مصطلحي « الفهرسة الموضوعية » و « التكشيف الموضوعي » ، حيث يدل الأول على الوحدات الوراقية الكاملة ، والثاني على أجزاء الوحدات ، تمييز مصطنع ومضلل وغير مطرد . فالعملية التي يتم بها التعبير عن المحتوى الموضوعي لأوعية المعلومات في مراصد البيانات المنشورة ، طباعة أو في شكل قابل للقراءة بواسطة الآلات ، غالباً ما تعرف وبشكل مطرد بالتكشيف الموضوعي ، سواء كان الأمر يتعلق بالأوعية الكاملة أو بأجزاء منها . وهكذا ، فإن الكشف الموضوعي لمرصد *Chemical Abstracts* مثلاً يمكن أن يشير إلى الكتب الكاملة وتقارير البحوث الكاملة فضلاً عن أجزاء الأوعية (كفصول الكتب وبحوث المؤتمرات ومقالات الدوريات) . ومن ناحية أخرى ، يمكن للمكتبات أن تفضل التعريف بأجزاء الكتب (الفصول أو البحوث) في الفهرس ، ويسمى ذلك عادة بالفهرسة التحليلية ، وعندما يتعلق الأمر بالمحتوى الموضوعي فإن هذا النشاط يمكن أن يسمى بالفهرسة الموضوعية التحليلية .

وإذا نظرنا في المصطلح « تصنيف » فإن الموقف قد يبدو أكثر غموضاً؛ فالمكتبيون يميلون إلى استعمال هذا المصطلح للدلالة على استخدام أرقام التصنيف (المستقاة من بعض خطط التصنيف ، كتصنيف ديوي العشري أو التصنيف العشري العالمي ، أو تصنيف مكتبة الكونجرس لأوعية المعلومات ، وخاصة لأغراض ترتيب هذه الأوعية على أرفف المكتبات ، أو في وحدات الحفظ ... إلخ) . أما الفهرس الموضوعي للمكتبة فإما أن يكون هجائياً (الفهرس الموضوعي الهجائي أو الفهرس المعجمي) أو مرتباً وفقاً للتسلسل في إحدى خطط التصنيف (الفهرس المصنف) . فإذا تناول أحد المكتبيين مثلاً كتاباً وقرر أنه يدور حول « الطيور » ، حينئذ يمكن أن يعطي هذا الكتاب رأس الموضوع « الطيور » ، كما أنه بدلاً من ذلك يمكن أن يعطيه رقم التصنيف ٥٩٨.٢ . ويمكن لكثيرين تسمية العملية الأولى بالفهرسة الموضوعية والثانية بالتصنيف ، وهذا تمييز ساذج إلى أبعد الحدود . ويحدث مزيد من الخلط عندما نتبين أن الكشف الموضوعي يمكن أن ينطوي على استخدام إحدى خطط التصنيف ، وأنه من الممكن للكشاف الموضوعي المطبوع أن يتبع تسلسل إحدى خطط التصنيف ، ولا معنى على الإطلاق لمحاولات التمييز المصطلحي هذه حيث لا تسهم إلا في زيادة الخلط . وحقيقة الأمر أن التصنيف ، بأوسع معانيه ، يدخل في نسيج جميع الأنشطة المتصلة باختزان المعلومات واسترجاعها . ويرجع جانب من الخلط المصطلحي إلى العجز عن التمييز بين تحليل المحتوى أو تحليل المفاهيم ، ومراحل الترجمة في الكشف .

وإذا حدث على سبيل المثال أن تناول أحد اختصاصيي المعلومات إحدى الوثائق وقرر أنها تدور حول الإنسان الآلي ، فإن النشاط الفكري الذي ينطوي عليه القرار هو نفس النشاط ، سواء كانت الوثيقة التي تناولها اختصاصي المعلومات كتاباً ، أو فصلاً ، أو جزءاً من كتاب ، أو دورية ، أو مقالة في دورية ، أو أعمال مؤتمر ، أو بحثاً في مؤتمر ، أو أياً كان نوعها . وقد صنف اختصاصي المعلومات الوثيقة ، أي وضعها في الفئة الموضوعية « للوثائق التي تتناول الإنسان الآلي » .

وتنطوي عملية الترجمة على التعبير عن ناتج التحليل الموضوعي بمصطلح واحد أو عدة مصطلحات مستقاة من لغة ما . وما المصطلح الذي يخصص أو يعين للوثيقة إلا وسيمة تدل على فئة بعينها من الوثائق . ويمكن لهذه الوسيمة أن تكون الكلمة الإنجليزية ROBOTS ، المستقاة من أحد المكانز أو إحدى قوائم رءوس الموضوعات ،

أو من الوثيقة نفسها ، كما يمكن أن تكون الكلمة المقابلة في لغة أخرى ، أو مجموعة الرموز الرقمية ٦٢٩,٨٩٢ المستقاة من إحدى خطط التصنيف .

وعملية تحديد ما تدور حوله الوثيقة ، وتعيين وسيمة للوثيقة تعبر عن هذا القرار ، واحدة في جوهرها ، سواء كانت الوسيمة التي تم تعيينها مستقاة من خطة للتصنيف ، أو من مكنز ، أو من قائمة لرءوس الموضوعات ، وسواء كانت الوثيقة مادة ورقية كاملة ، أو جزءاً من مادة ، وسواء كانت الوسيمة سوف ترتب فيما بعد هجائياً أو في أي تسلسل آخر (أو لن ترتب على الإطلاق) ، وسواء كان الهدف من العملية هو تنظيم الأوعية على الأرفف ، أو التسجيلات في الفهارس ، أو في كشافات مطبوعة ، أو في مرصد بيانات قابلة للقراءة بواسطة الآلات .

ويدل تصنيف الوثائق ، في مجال اختزان المعلومات واسترجاعها على تكوين فئات من الوثائق على أساس محتواها الموضوعي ، وما المكانز وقوائم رءوس الموضوعات وخطط التصنيف الوراقية ، في الأساس ، سوى قوائم بالوسيمات التي يمكن بواسطتها التحقق من هذه الفئات وترتيبها . وتنطوي عملية البحث عن المعلومات على تحديد أي الفئات يمكن الرجوع إليها في الكشاف المطبوع أو الفهرس البطاقي أو مرصد البيانات القابل للقراءة بواسطة الآلات . ويمكن للبحث أن ينطوي على النظر في فئة واحدة (ككل ما ورد تحت المدخل ROBOTS مثلاً) كما يمكن أن ينطوي على التوافق بين الفئات (كالوثائق التي وردت تحت ROBOTS وتحت Artificial Intelligence أيضاً ، مثلاً) . ويتوقف عدد التوافق المحتملة ، أو مدى سهولة التوفيق بين الفئات المختلفة ، إلى حد بعيد ، على شكل الأداة المستخدمة للبحث ، وعلى ما إذا كانت مطبوعة أو إلكترونية على وجه الخصوص .

وقصارى القول ، فإن التكشيف الموضوعي مطابق نظرياً تماماً للفهرسة الموضوعية . وجوهر النشاط هنا هو التصنيف الموضوعي ، أي تكوين فئات من الوثائق على أساس محتواها الموضوعي . ويدل مصطلحاً « التكشيف الموضوعي » و« التكشيف » في هذا الكتاب ، على جميع أنشطة التصنيف الموضوعي . وتحظى مشكلات استرجاع المعلومات بالمعالجة من منظور عريض ، ولا وجه للتمييز نظرياً بين الاسترجاع الموضوعي عن طريق فهارس المكتبات ، والاسترجاع الموضوعي عن طريق

الأشكال الأخرى لمراصد البيانات ، كالكشافات المنشورة بالشكل المطبوع ، أو على الخط المباشر ، أو على أسطوانات ضوئية مكتنزة مثلاً ، على الرغم من أننا نسوق الأمثلة من جميع أنواع الأدوات للدلالة على أنها لا تتساوى جميعاً في الإمكانيات .

لقد مهدت هذه النظرة العامة الشاملة للمكونات الأساسية لنشاط استرجاع المعلومات ، والعوامل الأساسية المؤثرة في أداء نظم استرجاع المعلومات . وسوف نتناول هذه الموضوعات بشيء من التفصيل فيما يلي من فصول .

الفصل الثاني

صناعة مرصد البيانات

لم يحظ التاريخ المبكر لاسترجاع المعلومات بالحاسبات الآلية ، بالتوثيق الجيد . وواقع الأمر أننا لا نعرف على وجه اليقين أي النظم يمكن النظر إليه باعتباره أول نظام لاسترجاع المعلومات يعتمد على الحاسب الآلي . ومن بين أقدم النظم المعتمدة فعلاً على الحاسبات الآلية ، تلك التي أنشئت في مختبرات مدفعية الأسطول Naval Ordnance Laboratory في سلفر سبرنج بولاية ماريلاند عام ١٩٥٩م ، والنظام الذي أعدته جامعة وسترن ريزيرف للجمعية الأمريكية للمعادن حوالي عام ١٩٦٠م .

ومن الممكن القول باطمئنان إن النظم الرئيسية الرائدة لاسترجاع المعلومات في الولايات المتحدة الأمريكية ، نشأت في الحكومة الاتحادية في مطلع ستينيات القرن العشرين . وربما كانت أهم هذه النظم هي تلك التي أنشأتها وكالة المعلومات التقنية للقوات المسلحة (التي سميت فيما بعد مركز توثيق الدفاع وتعرف الآن باسم مركز المعلومات التقنية للدفاع) والإدارة القومية للملاحة الجوية والفضاء عام ١٩٦٢م ، والمكتبة القومية للطب التي انطلقت خدماتها المعروفة بالمدرز MEDLARS عام ١٩٦٣م . وينبغي النظر إلى هذه الأجهزة باعتبارها الرائدة في التجهيز الوراقى على نطاق واسع بواسطة الحاسبات الآلية ، على الرغم من أن كثيراً من الهيئات الأخرى قد سارت على دربها .

ولقد كان من الصعب في العقد السابع من القرن العشرين تبرير تخصيص نظام

إلكتروني لأغراض البحث الراجع للتسجيلات الوراقية فقط ، من الناحية الاقتصادية . وقد استمدت معظم النظم الوراقية الضخمة مبرراتها من براعتها الفنية الفائقة ؛ فقد كانت تميل ، وما زالت في كثير من الحالات ، لأن تكون متعددة الأغراض ، تنتج العديد من المخرجات أو الخدمات اعتماداً على المدخلات نفسها . وقد تطور العديد من هذه النظم كنتاج طبيعي لعملية النشر الآلي ، والحاجة إلى معالجة التسجيلات الوراقية في شكل قابل للقراءة بواسطة الآلات ؛ لأغراض مراجعة الأخطاء ، والفرز ، وتوحيد أشكال الصيغ Formatting واستخدام الحاسب في التنضيد . وكان من الممكن استخدام الأشرطة القابلة للقراءة بواسطة الآلات الناتجة عن هذا النشاط في إنتاج المزيد من المطبوعات وتقديم المزيد من الخدمات . وكانت الخدمة الرئيسية التي تتيحها مرصد البيانات الإلكترونية هي خدمة البحث الراجع عند الطلب (خدمة البحث وفقاً للطلب) وإن كانت مرصد البيانات تستخدم أيضاً في أنشطة الإحاطة الجارية (البث الانتقائي للمعلومات SDI) .

نظم العمل على دفعات :

لقد كانت خصائص نظم الاسترجاع الإلكترونية العاملة في ستينيات القرن العشرين متشابهة إلى حد بعيد ؛ فقد كانت هذه النظم تعمل على دفعات خارج الخط المباشر Offline ، وتعتمد على الأشرطة المغنطة كوسط للاختزان كما كان البحث فيها يتم تسلسلياً بوجه عام ، حيث كانت استراتيجيات البحث تضاهي تسلسلياً مقابل تسجيلات الوثائق للحصول على وراقيات مطبوعة . وكانت عمليات البحث الراجع لمرصد البيانات كاملاً تتم مرة واحدة في العادة ، كما كان الناتج يتكون من جميع تسجيلات الوثائق التي يضمها النظام والتي تضاهي طلب البحث . أما بالنسبة للبث الانتقائي للمعلومات ، فكانت سمات البحث المختزنة تضاهي دورياً مقابل أحدث تسجيلات أضيفت إلى مرصد البيانات ، ثم ترسل النتائج إلى المشتركين .

وكانت نظم الاسترجاع الإلكترونية في ستينيات القرن العشرين تتمتع بكثير من المزايا مقارنة بما سبقها . وكانت هذه المزايا تشمل ما يلي :

١ - كان من الممكن عن طريق التجهيز على دفعات إجراء العديد من عمليات البحث في الوقت نفسه .

٢ - كان من الممكن، وبشكل اقتصادي، توفير العديد من المداخل أو نقاط الوصول إلى الوثائق .

٣ - كان من الممكن التعامل مع عمليات البحث المركبة التي تنطوي على أعداد كبيرة من المصطلحات في علاقات معقدة .

٤ - كان من الممكن تقديم المخرجات في شكل وراقيات، كما كان من الممكن إنتاج نوعيات جيدة من المطبوعات بربط نظام الاسترجاع بجهاز للتنضيد الضوئي . كما كان من الممكن أيضاً تقديم المخرجات مباشرة على مصغرات فيلمية (مخرجات الحاسب على مصغرات فيلمية COM) .

٥ - كان من الممكن تجميع البيانات الإدارية المتعلقة بالجوانب النوعية والكمية للإفادة من النظام، وعلى أساس منتظم وكناتج جانبي للعمليات العادية للنظام .

٦ - كان من الممكن الحصول على العديد من المخرجات والخدمات من المدخلات نفسها، كالكتشافات العامة المطبوعة، والكشافات المتخصصة، وعمليات البحث الراجع، وعمليات البحث لأغراض البث الانتقائي للمعلومات .

٧ - كان من الممكن بمجرد تسجيل مرصد البيانات في شكل قابل للقراءة بواسطة الآلات، استنساخه ببساطة وبتكلفة محدودة، كما كان من الممكن تداوله وببساطة في أي مكان، ومن ثم إمكان الإفادة منه في تقديم خدمات المعلومات من جانب مختلف مراكز المعلومات .

ورغم كثرة ما كانت تتمتع به نظم التجهيز على دفعات خارج الخط المباشر من مزايا، فقد كانت تكتنفها أيضاً بعض العيوب؛ فقد كانت هذه النظم في الأساس نظم فرصة البحث الواحدة، حيث كان يتعين على المسئول عن البحث أن يتدبر جميع مداخل البحث المحتملة مقدماً، وأن يصوغ استراتيجية البحث التي كان من الممكن عند مضاهاتها بمرصد البيانات، أن تسترجع جميع مفردات الإنتاج الفكري المناسب . وبعبارة أخرى، فإن هذه النظم كانت غير تفاعلية وغير إيعازية، كما أنها لم تكن تكفل أية إمكانية تصفح حقيقية .

ومن بين العيوب الجوهرية لنظم التجهيز على دفعات أيضاً تأخر الحصول على

نتائج البحث بشكل ملحوظ ؛ فلم يكن من الممكن الحصول على استجابة فورية . وكان الحصول على نتائج البحث يستغرق عدة ساعات في أفضل الظروف ، أما في أسوأ الظروف ، كما في حالة عمليات البحث التي يتم إجراؤها عن طريق أحد مراكز المعلومات الوطنية الكبرى ، فقد كان الحصول على النتائج يستغرق عدة أيام وربما عدة أسابيع .

وكان البحث في نظم التجهيز على دفعات « مفوضاً » بوجه عام ، أي أنه كان على من يحتاج إلى المعلومات أن يعهد بمهمة إعداد استراتيجية البحث لأحد اختصاصيي المعلومات ، حيث لم تكن الفرصة متاحة له لإجراء بحثه بنفسه . وليس من الضروري أن يكون البحث غير المفوض أفضل دائماً من البحث المفوض ، إلا أن عملية التفويض تتطلب قدراً من الحيلة والحذر .

فمما لا شك فيه أنه من الممكن للبحث أن يسفر عن نتائج سيئة ، إذا كان طالب الحصول على المعلومات عاجزاً ، في عملية التفويض ، عن التعبير بوضوح عما يبحث عنه ، أو ما إذا أساء اختصاصي المعلومات تفسير احتياجات المستفيد الحقيقية .

نظم الخط المباشر :

لقد أعقب نظم التجهيز على دفعات التي شهدتها ستينيات القرن العشرين نظم استرجاع تفاعلية على الخط المباشر في السبعينيات وما بعدها . وقد توافرت مقومات هذه النظم بفضل التطورات التي حدثت في العتاد والبرمجيات والاتصالات بعيدة المدى .

وفي نظم الخط المباشر يتم اختزان البيانات على أسطوانات ممغنطة . وعادة ما يتكون النظام من ملف خطي أو تسلسلي (يشتمل على جميع تسجيلات النظام كاملة) و ملف مصنف واحد أو أكثر (غالباً ما تسمى الملفات الكشفية) تستمد من الملف الخطي . ويوضح الشكل رقم (٦) مثلاً مبسطاً لعدد قليل من تسجيلات الملف الخطي ، مصحوباً بمثال لمداخل الملفات المصنفة . ويتكون كل مدخل من مداخل الملف المصنف من قيمة أو عنصر من عناصر تسجيلة مرصد البيانات الواردة في الملف الخطي (كاسم المؤلف أو إحدى الكلمات المفتاحية مثلاً) مصحوباً بعنصر رئيس متفرد (عادة ما يكون رقم القيد) يستخدم لاسترجاع التسجيلات التي يمكن أن ترد بها تلك القيمة أو ذلك العنصر . ويتم اختزان كل من الملفات الخطية والملفات المصنفة على أسطوانات

ممغنطة ، حيث يمكن الوصول إلى المعلومات عشوائياً ، ومن ثم إتاحة القدرة على إجراء عمليات البحث التفاعلية بشكل فوري .

Chan, Lois Mai
Pollard, Richard C.
Thesauri used in online databases
Greenwood Press:us
1988
United States
Language: English
Subject heading: Thesauri/Bibliography
Subject heading: Information systems/Directory
BLIB88009087
Monograph

ملف خطي

Instructional materials used in teaching cataloging
and classification
Chan, Lois Mai
Cataloging & Classification Quarterly 7:131-44 Summ '87
Language: English
Subject heading: Cataloging/Teaching
Subject heading: Surveys/Library science literature
Subject heading: Textbooks
BLIB87009368
Article

المؤلف

| | |
|---------------------|----------|
| Chan, Lois Mai | 88009087 |
| | 87009368 |
| Pollard, Richard C. | 88009087 |

ملفات

مصنفة

الموضوع

| | |
|------------------------------------|----------|
| Bibliography | 88009087 |
| Cataloging/Teaching | 87009368 |
| Directory | 88009087 |
| Information systems | 88009087 |
| Information systems/Directory | 88009087 |
| Library science literature | 87009368 |
| Surveys | 87009368 |
| Surveys/Library science literature | 87009368 |
| Teaching | 88009087 |
| Textbooks | 87009368 |
| Thesauri | 88009087 |
| Thesauri/Bibliography | 88009087 |

الشكل رقم (٦) : عينة للملفات المصنفة والتسلسلية ، مقتبسة من مرصد بيانات *Library Literature*

ولقد كان للتطورات التي شهدتها تقنيات الاتصالات بعيدة المدى أثرها الواضح في صناعة الاسترجاع على الخط المباشر ، وعلى الرغم من أن نظم الخط المباشر كانت متوافرة فعلاً في نهاية ستينيات وبداية سبعينيات القرن العشرين ، فإنها لم تستخدم على نطاق واسع ، وخاصة في المكتبات نظراً لما كان يستلزمه التعامل معها عبر خطوط الهاتف بعيدة المدى من تكلفة . وقد أدت شبكات التحويل الحزمي Packet - switching كتلتن TELENET وتايمنت TYMNET إلى خفض تكلفة الاتصالات بشكل ملحوظ ، نظراً لأن الشبكة لم تكن تعمل في خدمة أي مستفيد إلا عندما تكون الرسالة جاهزة للإرسال ، ولم تكن تحتفظ برابطة دائمة طوال الوقت الذي يستغرقه الاتصال . ولم يتحقق لسوق الاسترجاع على الخط المباشر الرواج إلا بعد دخول هذه الشبكات .

وتتمتع نظم الاسترجاع على الخط المباشر بجميع مزايا نظم التجهيز على دفعات ، ولكنها تتجنب جميع عيوبها الجوهرية . فهذه النظم إيعازية وتفاعلية وتتيح إمكانية التصفح ، وتكفل سرعة الاستجابة ، كما يمكن أن تستخدم لإجراء عمليات البحث بواسطة المستفيدين أنفسهم دون تفويض .

وكانت جميع النظم المبكرة للاسترجاع على الخط المباشر تقريباً تعمل كما في الشكل رقم (٥) ؛ فالوثائق الأولية يتم اقتناؤها من جانب إحدى المؤسسات (منتج مرصد البيانات) حيث يتم إعداد تسجيلات الوثائق التي تشتمل على الوصف الوراقى المناسب ، فضلاً عن المصطلحات الكشفية والمستخلصات في معظم الأحيان . وكانت هذه التسجيلات يتم إدخالها في شكل قابل للقراءة بواسطة الآلات ، واختزانها على أشرطة ممغنطة . كما كانت المعلومات المسجلة على الأشرطة يتم تجهيزها على أكثر من نحو لإنتاج الكشافات المطبوعة ، كما كان المنتج يحتفظ بها في بعض الأحيان لإجراء عمليات البحث على دفعات خارج الخط المباشر . كذلك كانت المعلومات المسجلة على أشرطة تحمّل على أسطوانات ممغنطة ، كما كانت التسجيلات يتم تجهيزها لإعداد الملفات الكشفية اللازمة للبحث التفاعلي على الخط المباشر . وكانت عمليات التجهيز هذه تتم إما بتأجير الأشرطة لمؤسسة أخرى (مورد مرصد البيانات) أو محلياً بواسطة منتج مرصد البيانات . وأياً كان المسئول عن التجهيز ، فإنه كان يتعين على المؤسسة التي تقوم بتجهيز مرصد البيانات للبحث على الخط المباشر ، أن توفر أيضاً برمجيات البحث التفاعلي المناسبة .

وقد مرت النظم التي تقدم البيانات الوراقية لفهارس المكتبات بمراحل التطور نفسها، إلا أنه بدلاً من أن يقوم المنتجون بتأجير مراصد البيانات للموردين ، تقوم المكتبات الأعضاء بإدخال التسجيلات في مرصد بيانات مركزي تملكه وتديره مؤسسة تعرف بالمرفق الوراقى . ثم تقوم مثل هذه المؤسسة ، كمركز الحاسب الآلى للمكتبات على الخط المباشر (OCLC) Online Computer Library Center ، مثلاً ، بتقديم الخدمات المناسبة والدعم التقني للمكتبات ، لكي تتلقى من مرصد البيانات المركزي نسخاً من التسجيلات المقابلة للأوعية التي تضمها مقتنياتها . ومن الممكن لهذه التسجيلات ، سواء كانت مطبوعة على الورق أو في شكل قابل للقراءة بواسطة الآلات ، أن تضاف إلى فهرس المكتبة .

وكانت المرافق الوراقية ، التي تعمل منذ البداية وفقاً لمواصفات الفهرسة المقروءة ألياً (مارك) Machine Readable Cataloging (MARC) ، لاختران تسجيلات الفهرسة وتجهيزها (الشكل رقم ٧) ، يتم التعامل معها في المراحل المبكرة على دفعات فقط ، لإنتاج بطاقات الفهارس للمكتبات الأعضاء . وقد حدث بعد ذلك في سبعينيات القرن العشرين أن أصبحت مراصد البيانات التي تكونت على هذا النحو ، متاح أيضاً على الخط المباشر ، حتى يمكن للمكتبات إجراء عمليات البحث وتعديل التسجيلات بشكل تفاعلي .

وعلى الرغم من أن التطور التقني لكل من المرافق الوراقية ومنتجي مراصد البيانات ومورديها كان مماثلاً تماماً ، فإن هناك أيضاً بعض الاختلافات الجوهرية بينها . فقد نشأت المرافق الوراقية نتيجة لاستخدام جميع المكتبات التي تنتج التسجيلات وتدخلها في مرصد البيانات المركزي ، المواصفات نفسها ، وهي الطبعة الثانية من قواعد الفهرسة الأنجلو - أمريكية *Anglo-American Cataloging Rules* أو AACR 2 ، لتحديد محتوى تسجيلات الوثائق ، ومواصفات مارك لتوحيد صيغ هذه التسجيلات وصيانتها ومعالجتها في شكل قابل للقراءة بواسطة الآلات . وفيما عدا بعض الاستثناءات المحدودة ، تتقبل فهارس المكتبات على الخط المباشر تسجيلات مارك وتعمل بها . ومن ناحية أخرى تبدو صناعة مراصد البيانات في معظم الأحيان بمنأى عن توحيد أساليب العمل ، حيث يختلف محتوى التسجيلات وبنيتها من مرصد إلى آخر بشكل ملحوظ . أضف إلى ذلك أنه على الرغم من تزايد تنوع الوثائق التي تغطيها كل

هذه النظم ، فإن المداخل التي تغطيها المرافق الوراقية والفهارس المتاحة على الخط المباشر عادة ما تكون الإشارات الوراقية المتصلة بالوثائق التي تضمها مقتنيات المكتبات. إلا أن صناعة مراصد البيانات لا تقدم البيانات الوراقية فحسب ، وإنما تقدم أيضاً أنواعاً أخرى من التسجيلات ، بما في ذلك البيانات الرقمية وكذلك النصوص الكاملة للوثائق .

ويتسم تتابع الأحداث الذي نقدمه في هذا القسم بالبساطة الشديدة ، كما أنه يسلك في الواقع أكثر من مسار رئيس واحد . ولم تعد الحدود الفاصلة بين ما يمكن أن تضطلع به مختلف المؤسسات من أدوار ، وأنواع ما تكفل الوصول إليه من معلومات ،

► NO HOLDINGS IN OCL - FOR HOLDINGS ENTER dh DEPRESS DISPLAY RECD SEND
OCLC: 3349989 Rec stat: n Entrd: 771108 Used: 790312 ¶
► Type: a Bib lvl: m Govt pub: Lang: eng Source: illus: a
Repr: Enc lvl: Cont pub: 0 Ctry: nyu Dat tp: ■ M/F/B: 10
Indx: 1 mod rec: Festschr: 0 Cont: b
Desc: i Int lvl: Dates: 1977. ¶
► 1 010 77-77941 ¶
► 2 040 DLC ꝥ DLC ¶
► 3 020 0525171940 : ꝥ \$17.95 ¶
► 4 050 0 GN31.2 ꝥ .L43 1977 ¶
► 5 082 573.2 ¶
► 6 090 ꝥ ¶
► 7 049 OCOC ¶
► 8 100 10 Leakey, Richard E. ¶
► 9 245 10 Origins : ꝥ what new discoveries reveal about the
emergence of our species and its possible future / ꝥ 25 Richard E. Leakey
and Roger Lewin. ¶
► 10 260 0 New York : ꝥ Dutton, ꝥ c1977. ¶
► 11 300 264 p. : ꝥ ill. (some col.) ; ꝥ 25 cm. ¶
► 12 504 Bibliography: p. 257. ¶
► 13 500 Includes index. ¶
► 14 650 0 Anthropology. ¶
► 15 650 0 Human evolution. ¶
► 16 700 10 Lewin, Roger, ꝥ joint author. ¶

لم تعد هذه الحدود الفاصلة واضحة كما كانت من قبل ، كما بدأت الساحة تتسع لأنواع أخرى كثيرة من مراصد البيانات والمؤسسات الضالعة في المجال . ويقدم جانب كبير مما تبقى من هذا الفصل ما يدل على ثراء وتنوع هذه البيئة دائمة التغير .

الاتجاهات السائدة في الصناعة :

تدل جميع المؤشرات على أن صناعة الخط المباشر في نمو مستمر . ومن الممكن الحصول على فكرة جيدة عن حجمها وتنوعها في الوقت الراهن ، بتصفح أحد الأدلة الحديثة لمراصد البيانات مثل :

- *Computer - Readable Databases* (Marcaccio, Adams and Williams , 1990) .

- *Directory of Online Databases* (Cuadra Associates, 1992) و

وعلى الرغم من دينامية هذه السوق ، حيث لا تعمر مراصد البيانات والموردون طويلاً في بعض الأحيان ، فإن الاتجاه العام يميل إلى الزيادة المطردة .

وتسهم في هذا النمو عوامل عدة . ويسجل وليامز (1988 , 1992) Williams أن عدد مراصد البيانات المتاحة على الخط المباشر وعلى دفعات تضاعف من ٣٠٠ إلى ٦٠٠ بين عامي ١٩٧٥ م و ١٩٨١ م ، ثم ارتفع إلى خمسة أمثاله من ٦٠٠ إلى ٣٠٠٠ بين عامي ١٩٨١ م و ١٩٨٥ م . وقد بلغ عدد عمليات البحث التي أجريت عام ١٩٧٥ م مليون عملية ، وفي عام ١٩٨٥ م تم إجراء ١٥ مليون عملية . كذلك ارتفع عدد تسجيلات مراصد البيانات من ٥٢ مليون تسجيلية عام ١٩٧٥ م إلى ١,٦٨ بليون عام ١٩٨٥ م (Williams , 1988 , 1980 A) . ويسجل وليامز أيضاً (1992) Williams أنه في عام ١٩٩١ م أجريت ٣٤,٥ مليون عملية بحث في ٧٦٣٧ مرصداً للبيانات ، تشتمل على أربعة بلايين تسجيلية .

هذا ، وتشتمل إصدارة يوليو ١٩٩٢ م من دليل *Directory of Online Databases* Cuadra Associates, 1992 ، الذي يغطي فقط المصادر المتاحة على الخط المباشر ، على أكثر من ٥٣٠٠ مرصد للبيانات ، يقوم بإنتاجها ٢١٥٨ منتجاً ، ويمكن البحث فيها من خلال ٧٣١ مورداً . وهذه المراصد وخدمات الاسترجاع على الخط المباشر متاحة في جميع أنحاء العالم . وتدويل صناعة الاسترجاع على الخط المباشر من الاتجاهات النامية التي

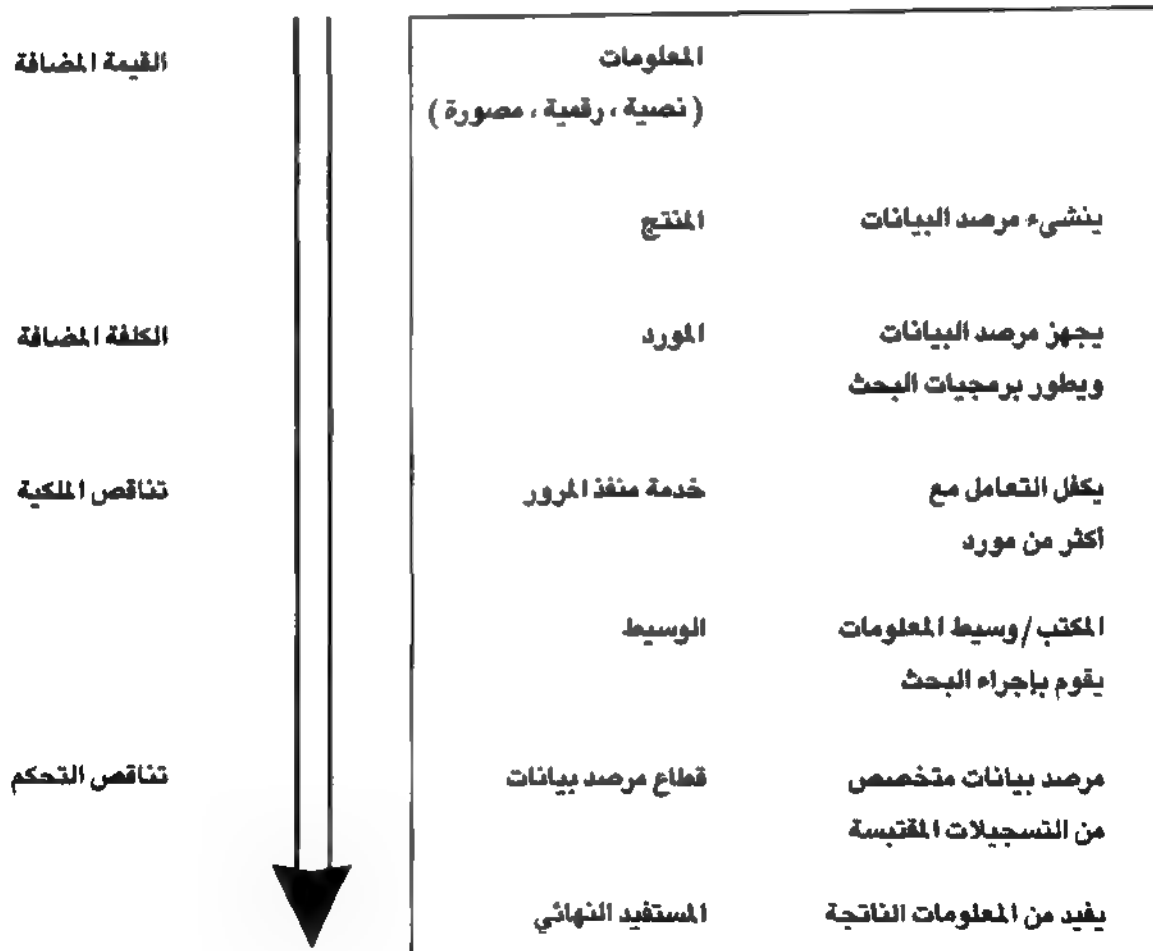
يمكن أن تستمر (Landau , 1988) ، نتيجة لتزايد أعداد الشركات العاملة في خارج الحدود ، فضلاً عن الاستثمارات الأجنبية التي تتم في كثير من المناطق بما في ذلك الولايات المتحدة الأمريكية .

وتسلك المرافق الوراقية والفهارس المتاحة على الخط المباشر مساراً مماثلاً في تطورها ؛ ففي عام ١٩٧٦م لم يكن هناك سوى مليونين فقط من تسجيلات الفهرسة في مرصد البيانات الإلكتروني الذي يربطه مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC أضخم مرفق وراقي . وفي عام ١٩٩١م ارتفع الرقم إلى أكثر من ٢٣ مليون تسجيلة (OCLC, Inc. 1991) ولم يكن هناك في سبعينيات القرن العشرين سوى ما لا يتجاوز أصابع اليد الواحدة من المؤسسات الأكاديمية التي توافرت لها البنى الأساسية ، الاقتصادية والتقنية الثرية الكافية لتطوير ودعم الفهارس المتاحة على الخط المباشر ، أما الآن ، فإن التطورات التقنية ، وخاصة مع توافر برمجيات وعتاد الفهارس المتاحة على الخط المباشر ، بتكلفة يمكن تحملها ، قد جعلت من الممكن للمكتبات على اختلاف أنواعها وتباين أحجامها ، إنشاء وصيانة الفهارس على الخط المباشر . وهناك الآن ولا شك آلاف من هذه المكتبات .

ولا تتسم صناعة الخط المباشر بالضخامة فقط ، وإنما تتسم أيضاً بالتعقد البالغ . قارن الشكل رقم (٥) في الفصل الأول ، الذي يبين منتجي مراصد البيانات ومتعهديها والأدوار التي ينهضون بها ، بالشكل رقم (٨) الذي يضع هذه المؤسسات جنباً إلى جنب مع المؤسسات الأخرى العاملة في سلسلة الإفادة من مراصد البيانات (Williams, 1986) . وقد سبق أن ذكرنا دور كل من المنتج (إنشاء مراصد البيانات) والمتعهد (تجهيز هذه المراصد وجعلها قابلة للبحث على الخط المباشر) ، ثم ظهرت حديثاً خدمات بوابات المرور Gateway (حوالي منتصف ثمانينيات القرن العشرين) ، والتي أدخلت على وجه التحديد نتيجة لتزايد أعداد مراصد البيانات وخدمات الخط المباشر . وتتيح خدمات بوابات المرور هذه للمستخدمين إمكانية التعامل مع العديد من خدمات الخط المباشر النائية ، من خلال حاسب آلي واحد ، يمثل منفذاً رئيساً ، يكفل أيضاً في معظم الأحيان واجهة تعامل موحدة مع هذه الخدمات ، وهذا عنصر مفيد جداً ، حيث لا يجعل هناك حاجة إلى تعلم لغات التخاطب المختلفة اللازمة للتعامل مع كل نظام . ومن الممكن البحث في

مرصد البيانات بواسطة وسطاء (اختصاصيو المعلومات في المكتبات أو شركات سمسة المعلومات) وأحياناً ما تستخدم النتائج في إنشاء مرصد بيانات فرعي جديد، يسمى غالباً بمرصد البيانات « المحلي » يمكن حينئذ البحث فيه محلياً في المكتبة أو مركز المعلومات . وأخيراً يقوم المستفيد النهائي بتفسير المعلومات الناتجة والإفادة منها .

وعلى الرغم من وجود كل هؤلاء الأطراف الآن في صناعة المعلومات المعاصرة ، فإنهم قد لا يشاركون دائماً كلما أجريت عملية بحث . هذا بالإضافة إلى أنه من الممكن لنفس المؤسسة أو الفرد الاضطلاع بأكثر من دور واحد في هذه السلسلة ؛ فالمكتبة القومية للطب (NLM) على سبيل المثال ، تقوم بإنتاج مرصد البيانات الخاص بها والمعروف باسم المدلين MEDLINE ، وتتيح إمكانية التعامل معه على الخط المباشر على أساس تجاري في الوقت نفسه .



الشكل رقم (أ) : سلسلة الإفادة من مرصد البيانات ، عن : Williams (1986) يتصرف

ومع التحرك نزولاً في سلسلة الإفادة من مرصدا البيانات تتزايد القيمة نتيجة لتجهيز البيانات أو تنقيتها . وغالباً ما يسفر ذلك عن زيادة في التكلفة التي عادة ما يتحملها المستفيد . ومع كل خطوة في السلسلة تزداد احتمالات الوصول إلى المعلومات المختزنة ، إلا أن ذلك يتحقق على حساب تناقص القدرة على السيطرة أو التحكم . هذا بالإضافة إلى أن تزايد القيمة يزيد من مخاطر تشويه الرسالة الأصلية ، حيث يمكن إسقاط أو إهمال أو تغيير أو إضافة شيء له أهميته .

وليس هذا فحسب ، وإنما يمكن أيضاً لاتساع المسافة بين المعلومات المنشورة والمستفيد ، وتزايد أعداد المؤسسات التي تكفل سبل الوصول إلى هذه المعلومات ، أن يؤدي إلى صعوبة تحديد من يمكن أن يقدر احتياجات المستفيد ويعمل على تلبيتها ، فضلاً عن تحديد أي احتياجات المستفيد أجدى من غيرها ، وأمكن تلبيتها على أحسن وجه عن طريق ناتج معين أو خدمة معلومات معينة ، وما إذا كانت هي احتياجات المستفيد الوسيط أم المستفيد النهائي . كذلك يؤدي هذا الاتساع في المسافة والتزايد في أعداد المؤسسات إلى تزايد صعوبة الاختيار بالنسبة للمستفيد الذي يقيم البدائل المتاحة لسبل الوصول إلى مصادر المعلومات .

كذلك يمكن لسلسلة الإفادة من مرصدا البيانات أن تشمل بالوصف الأدوار التي تضطلع بها كل من المرافق الوراقية والمكتبات والفهارس المتاحة على الخط المباشر ؛ فمن الممكن النظر إلى المرافق باعتبارها من منتجي تسجيلات الفهرسة بصيغ مارك ومن موردي هذه التسجيلات في الوقت نفسه . أما الوسطاء في هذه الحالة فهم المكتبيون الذين يبحثون عن تسجيلات مارك ويعدلونها ويطلبون نسخاً منها من مرصدا البيانات . وحينئذ يصبح الفهرس المتاح على الخط المباشر مرصداً فرعياً فعلاً ، يتم تجهيزه وتكشيفه وتهيئته للبحث بالطرق التي تناسب المستفيدين النهائيين من المكتبة .

منتجات المعلومات وخدماتها :

إن أهم منتج معلوماتي هو مرصد البيانات الإلكتروني الذي يمكن وصفه تبعاً لمجاله الموضوعي ، أو الشكل المادي لمحتواه ، أو تبعاً لمختلف أوجه الإفادة منه .

ونظراً لأنه لا يمكن لأي مرصد للبيانات أن يشتمل على جميع معلومات العالم ، فإن

جميع مراصد البيانات عادة ما تكون مقيدة إلى حد ما في مجالها وحدودها . ومن أبرز القيود المجال الموضوعي . وكانت مراصد البيانات المبكرة تقتصر على المعلومات العلمية والتقنية التي تحتاجها الأجهزة الحكومية التي ترعاها . وعندما بدأ القطاع التجاري إنتاج مراصد البيانات في سبعينيات القرن العشرين ، اتسعت مجالات تغطية خدمات الكشف والاستخلاص لتشمل الهندسة والعلوم التطبيقية في البداية ، ثم العلوم الاجتماعية (بما في ذلك الاقتصاد وإدارة الأعمال) والفنون والإنسانيات في نهاية المطاف . وتشتمل كل مراصد البيانات هذه تقريباً على المعلومات التي تدخل أساساً في صميم اهتمام العلماء الباحثين ورجال الأعمال والمهندسين . إلا أنه قد تم فيما بعد إنشاء مراصد البيانات التي تركز على المواد التي تهتم المواطن العادي ؛ من معلومات المستهلك حول العديد من المنتجات المتنوعة إلى المعلومات العملية حول رعاية الطفل والعقاقير ، وما شابه ذلك من الاهتمامات اليومية (Williams, 1985) .

هذا ومن الممكن أيضاً تقسيم مراصد البيانات تبعاً لشكل ما تشتمل عليه من بيانات، وتبعاً لأوجه الاستفادة المحتملة منها . ويقدم الإنتاج الفكري عدداً من التقسيمات الثنائية :

١ - مراصد البيانات النصية في مقابل المراصد الرقمية أو التصويرية : فهناك مراصد البيانات التي تشتمل أساساً على البيانات النصية ، ومراصد البيانات التي تشتمل على بعض الأشكال الأخرى للبيانات ، كالبيانات الرقمية والبيانات التخطيطية أو التصويرية (Williams, 1985) .

٢ - مراصد البيانات الوراقية في مقابل غير الوراقية : فمراصد البيانات التي تشتمل على إشارات وراقية للإنتاج الفكري تدخل في فئة ، بينما تدخل جميع مراصد البيانات الأخرى في فئة أخرى تشمل الأدلة ، والنصوص الكاملة ، ومراصد البيانات الرقمية (Borgman, Moghdam and Corbett, 1984) .

٣ - مراصد البيانات الإرشادية في مقابل مراصد البيانات المصدرية : فمراصد البيانات المرجعية هي المراصد الوراقية فضلاً عن مراصد البيانات التي تشتمل على مداخل من مصادر كأدلة المؤسسات والأفراد والبرمجيات . أما مراصد البيانات

المصدرية فهي جميع مراصد البيانات الأخرى والتي تشتمل على بيانات أولية
كالمعلومات الرقمية والنصوص الكاملة للوثائق (Harter, 1986) .

وهكذا ، يمكن وضع مراصد البيانات الإلكترونية في واحدة من خمس فئات مختلفة
تبعاً لمحتواها وأهدافها ومجالاتها :

١ - مراصد البيانات الوراقية / المرجعية / النصية التي تشتمل على إشارات وراقية
للإنتاج الفكري الأولي . وتستخدم هذه المراصد عادة لإجراء عمليات البحث
الراجع فضلاً عن البحث لأغراض البث الانتقائي للمعلومات ، لدعم جهود البحث
والأنشطة العلمية . وأحياناً ما تتحدد مجالات هذه المراصد موضوعياً كما يمكن أن
تتحدد تبعاً لأنواع الوثائق التي تعرف بها (كتقارير البحوث أو الكتب أحادية
الموضوع أو مقالات الدوريات مثلاً) ويوضح الشكل رقم (٩ أ) نموذجاً
لتسجيلية من مرصد بيانات *America ; History and Life* الذي يشتمل على الإشارات
الوراقية للإنتاج الفكري بكل أنواعه حول تاريخ الولايات المتحدة الأمريكية وكندا
والدراسات الإقليمية الخاصة بهما وشئونهما المعاصرة ، منذ عام ١٩٦٤ م .

٢ - مراصد البيانات الوراقية / الإرشادية / النصية ، التي تشتمل على معلومات
حول البشر أو الشركات أو مستودعات البحث أو الوسائط كالبرمجيات والمواد
السمعية البصرية . ولا تستخدم هذه المراصد للاستدلال على أوعية المعلومات ، وإنما
للإجابة عن الأسئلة المتعلقة بالمصادر غير المطبوعة ، وغالباً ما تستخدم للرد على
الاستفسارات المرجعية السريعة . ويوضح الشكل رقم (٩ ب) نموذج تسجيلية من
مرصد بيانات إرشادي وهو *American Men and Women of Science* ويشتمل على
تراجم ١٣٠ ٠٠٠ عالم أمريكي وكندي في العلوم الفيزيائية والبيولوجية .

٣ - مراصد البيانات غير الوراقية / المصدرية / النصية ، وهي مراصد بيانات
النصوص الكاملة التي تضم نصوص الوثائق الأصلية في شكل قابل للقراءة
بواسطة الآلات . ويمكن لهذه الوثائق أن تكون مقالات دوريات ، أو صحف أو
نشرات إخبارية ، أو موسوعات أو معاجم ، أو غير ذلك من أنواع الكتب المرجعية .
وتستخدم مراصد بيانات النصوص الكاملة للإجابة عن أسئلة الحقائق فضلاً عن
استرجاع الإشارات الوراقية للإنتاج الفكري ، وهي في هذه الحالة الأخيرة تؤدي

وظيفة المراقب الوراقية نفسها. ويشتمل الشكل رقم (١٠) على تسجيلية نص كامل من مجلة *Harvard Business Review* ، حيث يشتمل ذلك الملف على النصوص الكاملة لمقالات الدورية فضلاً عن مرصد البيانات الوراقية الخاص بهذه المقالات من عام ١٩٧٦م فصاعداً.

A: America: History and Life (Bibliographic/Reference)

950008 26-8

AMERICAN HOMESTEADERS AND THE CANADIAN PRAIRIES, 1899 AND 1909.

Percy, Michael B ; Woroby, Tamara

Explorations in Economic History 1987 24(1): 77-100.

NOTE: Based on published census records and other public documents; 2 fig., 3 tables, 12 notes, ref., appendix.

DOCUMENT TYPE: ARTICLE

ABSTRACT: The out-migration of American homesteaders to the Canadian prairies is best explained by human-capital investments in wheat farming in 1899 and 1909 and by the techniques of dry farming in 1899. Canadian promotional expenditures also contributed to higher rates of out-migration. High tenancy rate and low agricultural wages in the United States were not important contributors to out-migration. (P. J. Coleman)

DESCRIPTORS: Prairie Provinces ; USA ; Homesteading and Homesteaders ; Agriculture ; 1894-1913 ; Migration

HISTORICAL PERIOD: 1890D 1900D 1910D 1800H 1900H

HISTORICAL PERIOD (Starting): 1894

HISTORICAL PERIOD (Ending): 1913

B: American Men and Women ■ Science (Bibliographic/Referral)

0010037

Brown, Donald D

DISCIPLINE: BIOLOGY, GENERAL (00200207)

SUBJECT SPECIALTY: DEVELOPMENTAL BIOLOGY

BORN: Cincinnati, Ohio, Dec 30, 31 **MARRIED:** 57 **NO. OF CHILDREN:** 3

EDUCATION: Univ Chicago MS & MD 56

HONORARY DEGREES: DSci Univ Chicago 76 Univ Maryland 83

PROFESSIONAL EXPERIENCE: Intern Charity Hosp New Orleans La 56-57; res assoc biochem NIMH 57-59; spec fel Pasteur Inst Paris 59-60; spec fel 60-61 **MEM STAFF BIOCHEM CARNEGIE INST DEPT EMBRYOL** 61 to present, **DIR DEPT EMBRYOL** 76 to present

CONCURRENT POSITIONS: Prof Johns Hopkins Univ 69 ■ present

MEMBERSHIPS: Nat Acad Sci; Am Soc Biol Chem; Soc Develop Biol; Am Acad Arts & Sci; Am Soc Cell Biol (pres 92); Am Philos Soc

HONORS AND AWARDS: US Steel Found Award Molecular Biol Nat Acad Sci 73; V

D Mattia Lectr Roche Inst Molecular Biol 75; Boris Pregel Award NY Acad

Sci 77; Ross ■ Harrison Prize Int Soc Develop Biologists 81; Feodor

Lynen Medal 87

RESEARCH: Control of genes during development; isolation of genes

ADDRESS: Dept Embryol, Carnegie Inst Washington 115 W University Pkwy Baltimore , MD 21210

الشكل رقم (٩) : نماذج تسجيلات مرصد بيانات وراقية .

118565

825040

****COMPLETE TEXT AVAILABLE****

How Global Companies Win Out

Hout, Thomas - Boston Consulting Group ; Porter, Michael E. - Harvard Univ. Graduate School of Business Administration ; Rudden, Eileen - Boston Consulting Group

HARVARD BUSINESS REVIEW, Sep/Oct 1982, p. 98

TEXT:

Hold that obituary on American manufacturers. Some not only refuse to die but even dominate their businesses worldwide. At the same time Ford struggles to keep up with Toyota, Caterpillar thrives in competition with another Japanese powerhouse, Komatsu. Though Zenith has been hurt in consumer electronics, Hewlett-Packard and Tektronix together profitably control 50% of the world's industrial test and measurement instrument market. American forklift truck producers may retreat under Japanese pressure, but two U.S. chemical companies—Du Pont and Dow—dramatically outperform their competitors.

How do these American producers hold and even increase profitability against international competitors? By forging integrated, global strategies to exploit their potential; and by having a long-term outlook, investing aggressively, and managing factories carefully.

The main reason is that today's international competition in many industries is very different from what it has been. To succeed, an international company may need to change from a multidomestic competitor, which allows individual subsidiaries to compete independently in different domestic markets, to a global organization, which pits its entire worldwide system of product and market position against the competition. (For a more complete discussion of this distinction, see the ruled insert.)

The global company—whatever its nationality—tries to control leverage points, from cross-national production scale economies to the foreign competitors' sources of cash flow. By taking unconventional action, such as lowering prices of an important product or in key markets, the company makes the competitor's response more expensive and difficult. Its main objective is to improve its own effectiveness while eroding that of its competitors.

Not all companies can or should forge a global strategy. While the rewards of competing globally are great, so are the risks. Major policy and operating changes are required. Competing globally demands a number of unconventional approaches to managing a multinational business to sometimes allow:

Major investment projects with zero or even negative ROI. Financial performance targets that vary widely among foreign subsidiaries. Product lines deliberately overdesigned or underpriced in some markets. A view of country-by-country market positions as interdependent and not as independent elements of a worldwide portfolio to be increased or decreased depending on profitability. Construction of production facilities in both high and low labor-cost countries.

Not all international businesses lend themselves to global competition. Many are multidomestic in nature and are likely to remain so, competing on a domestic-market-by-domestic-market basis. Typically these businesses have products that differ greatly among country markets and have high transportation costs, or their industries lack sufficient scale economies to yield the global competitors a significant competitive edge.

Before entering the global arena, you must first decide whether your company's industry has the right characteristics to favor a global competitor. A careful examination of the economies of the business will highlight its ripeness for global competition. Simply put, the potential for global competition is greatest when significant benefits are gained from worldwide volume—in terms of either reduced unit costs or superior reputation or service—and are greater than the additional costs of serving that volume.

SEE RECORD 118564 FOR A CONTINUATION OF THIS TEXT.

٤ - مرصد البيانات غير الوراقية / المصدرية / الرقمية ، وهي مرصد رقمية تتوافر في العلوم وإدارة الأعمال والاقتصاد بوجه خاص . وهذه لا تستخدم لإجراء عمليات البحث الموضوعي المعقدة ، وإنما لاسترجاع معطيات الحقائق الرقمية بطريقة مباشرة عادة دون أية تعقيدات . وتكمن القوة الحقيقية لمرصد البيانات الرقمية في قدراتها التحليلية وإمكانات العرض ؛ فمن الممكن أن تستخدم ، على سبيل المثال ، في إجراء التحليلات الإحصائية المتطورة ، وفي إعداد التنبؤات المستقبلية ، وفي التحقق من المركبات الكيميائية ، كما أنها يمكن أن تعرض المعطيات أو الأرقام الخام ، كما يمكن أن تترجم هذه الأرقام وتعبّر عنها بمصورات أو رسومات بيانية (Berger and Wagner, 1982) ويشتمل الشكل رقم (١١) على نموذج بيانات سلسلة زمنية مقتطف من « *PTS U.S. Time Series* » ، ويشتمل على بيانات إحصائية لمدة تتجاوز عشرين عاماً ، فضلاً عن حساب معدل النمو اعتماداً على بيانات تغطي جميع جوانب الاقتصاد الوطني ، والسكان ، والصناعة والمالية ، وغير ذلك من أنشطة إدارة الأعمال .

0126913 USDA RS 92/04/00 P32 United States
rice. land use. yr begin 8/1.

| YEAR | YEAR |
|------|------|
| 1979 | 2.89 |
| 1980 | 3.38 |
| 1981 | 3.83 |
| 1982 | 3.29 |
| 1983 | 2.19 |
| 1984 | 2.83 |
| 1985 | 2.51 |
| 1986 | 2.38 |
| 1987 | 2.36 |
| 1988 | 2.93 |
| 1989 | 2.73 |
| 1990 | 2.90 |
| 1991 | 2.86 |

GROWTH RATE= - 1.8%

CC=1USA PC=0112000 EC=411

الشكل رقم (١١) : نموذج تسجيلية من مرصد بيانات : *U.S. Time Series*

٥ - مراصد البيانات غير الوراقية / المصدريّة / التصويرية ، التي تتكون من الصور الفعلية ، بالإضافة إلى النصوص أو الأرقام عادة . واسترجاع المعلومات التصويرية مجال بكر نسبياً لم يستكشف بعد ، على الرغم من الاهتمام بأوجه الاختلاف الجوهرية بين تكشيف واسترجاع النصوص وتكشيف الصور واسترجاعها . ويشتمل الشكل رقم (١٢) على نموذج من مرصد بيانات - TRADEMARKSCAN FEDERAL الذي يشتمل على جميع العلامات التجارية المسجلة السارية والمعلقة ، في إدارة براءات الاختراع والعلامات التجارية بالولايات المتحدة الأمريكية . ويستمد هذا المرصد جانباً كبيراً من قدرته الاسترجاعية من حقوله النصية ، إلا أنه من المزمع ، في المستقبل ، أن تكون عناصر التصميم الفعلي قابلة للبحث بشكل مباشر ؛ كالبحث مثلاً عن جميع العلامات التجارية السارية التي يشكل النسر جزءاً من تصميمها (Thompson, 1989) .

وبينما كان لجميع مراصد البيانات الإلكترونية تقريباً في البداية نظائرها المطبوعة ، فقد تطورت صناعة المعلومات إلى الحد الذي أصبحت فيه تقتصر كثير من مراصد البيانات على الشكل القابل للقراءة بواسطة الآلات فقط . وربما يكون هذا الاتجاه مرتبطاً بالتحول من النظم المعتمدة على الورق أساساً إلى نظم الاسترجاع الإلكترونية (Lancaster, 1978) ويصدق ذلك بوجه خاص بالنسبة لمراصد البيانات الرقمية ، حيث ينطوي جزء من عملية الاسترجاع على معالجة البيانات وتحليلها ، ومن الواضح أن ذلك لا يمكن أن يتحقق في الشكل المطبوع . إلا أن هناك كثيراً من مراصد البيانات الوراقية ومراصد بيانات النصوص الكاملة ، التي لا تتوافر إلا في الشكل الإلكتروني فقط ، نظراً لأن منتجيها لا يتوقعون لها سوقاً واسعة ، نتيجة لمحتواها الموضوعي المغرق في التخصص .

هذا ، وقد بدأ إنتاج مراصد البيانات على أسطوانات ضوئية في ثمانينيات القرن العشرين ، ومن الممكن الآن الحصول على كثير من مراصد البيانات على أسطوانات ضوئية مكنزة ، بما في ذلك مرصد بيانات إرك *Educational Resources Information Center* (ERIC) ومرصد بيانات مدلاين *MEDLINE* (الوراقين) ودليل المكتبات الأمريكية *American Library Directory* (الإرشادي) وموسوعة *Grolier's Electronic Encyclopedia* والموسوعة الأمريكية الأكاديمية *Academic American Encyclopedia*

(النص الكامل) و *American Profile* (الرقمي) . ويسجل نكولز وفان دن إلشاوت (1990) Nicholls and Van Den Elshout رغم ذلك ، أنه قد أصبح من الصعب تقسيم المواد المسجلة على أسطوانات ضوئية مكتنزة إلى فئات ، بشكل دقيق محكم نظراً لوجود

04183916 DIALOG File 226: TRADEMARKSCAN (r)-Federal
 DESIGN ONLY
 INTL CLASS: 5 (Pharmaceuticals)
 4.2 (Wiscellaneous Service Waris)
 U.S. CLASS: 44 (Dental, Medical & Surgical Appliances)
 100 (Miscellaneous Service Marks)
 T&T U.S. CLASS: 18(Medicines & Pharmaceutical Preparations)
 STATUS: Pending; Non-Final Action - Mailed
 GOODS/SERVICES: (INT. CL. 5) VETERINARY PRODUCTS (INT. CL.42)
 SERVICES, MEDICAL INFORMATION, AND INFORMATION SERVICES
 TO PET OWNERS AND VETERINARIANS
 SERIAL NO.: 74-183,916
 FIRST USE: June 14, 1991 (Intl Class 5)
 June 14, 1991 (Intl Class 42)
 FIRST COMMERCE: June 14, 1991 (Intl Class 5)
 June 14, 1991 (Intl Class 42)
 FILED: July 11, 1991
 ORIGINAL APPLICANT: PROFESSIONAL COMMUNICATIONS GROUP,
 INC. (New York Corporation), 40 SANDRINGHAM RD., P.O.BOX 10515,
 ROCHESTER, NY (New York), 16410, USA (United States of America)
 FILING CORRESPONDENT: PROFESSIONAL COMMUNICATIONS
 GROUP, INC., 40 SANDRINGHAM RE., P.O.BOX 10515, ROCHESTER,
 NY 14610
 DESIGN CODES:
 03 (ANIMALS)
 0301 (CATS, DOGS, WOLVES, FOXES, BEARS)
 030104 (DOMESTIC CATS)
 030108 (OTHER DOMESTIC DOGS)
 030124 (STYLIZED ANIMALS IN THIS DIVISION (0301))
 030126 (COSTUMED ANIMALS IN THIS DIVISION (0301) & THOSE
 WITH HUMAN ATTRIBUTES)
 02 (HUMAN BEINGS)
 0207 (HUMAN DRESS, ATTIRE OR APPEARANCE)
 020792 (OTHER MALE ANIMAL ATTIRE)
 10 (TOBACCO, SMOKERS' MATERIALS, FANS, TOILET
 ARTICLES, MEDICAL DEVICES & APPARATUS, & TABLETS,
 CAPSULES OR POWDERS)
 1007 (MEDICAL DEVICES & APPARATUS)
 100704 (BANDAGES, CASTS, SLINGS)

الشكل رقم (١٢) : نموذج تسجيلية من ، TRADEMARKSCAN - FEDERAL

أعداد متزايدة من أنواع المنتجات أو المحتويات على الأسطوانة نفسها. ومن أبرز الأمثلة على ذلك مرصد بيانات *OncoDisk* ، وهو عبارة عن مصدر مرجعي كامل يشتمل على إشارات وراقية من كشف *CancerLit* ، فضلاً عن تعليمات التعامل مع *PDQ (Phy- sician's Data Query)* وهو مجموعة من أدلة الأطباء والمؤسسات الطبية ، بالإضافة إلى النصوص الكاملة لعدد من الكتب الدراسية الأساسية . وفي عام ١٩٩٢ م كان هناك أكثر من ١٣٠٠ ناتج على أسطوانات ضوئية مكتنزة ، متاحة تجارياً ، ومازال هذا الشكل من الوسائط يمثل أحد القطاعات سريعة النمو في صناعة المعلومات .

متعهدو مرصد البيانات :

المتعهدون هم تلك المؤسسات التي تقوم بتجهيز مرصد البيانات وتهيئتها للبحث بواسطة برمجيات تفاعلية . ولقد كانت هناك زيادة مطردة في عدد المتعهدين أو الموردين ، من ٥٩ متعهداً عام ١٩٧٩ م إلى أكثر من ٧٠٠ عام ١٩٩٢ م . ويقدم بعض هؤلاء المتعهدين البنية الأساسية أو المقومات التقنية الأساسية اللازمة للبحث في مرصد البيانات المتاحة من خلالهم فقط ، والبعض الآخر يقوم بإنتاج مرصد البيانات أيضاً . ومن أكبر متعهدي مرصد البيانات *BRS Information Technologies* و *ORBIT Search Service* (كلاهما فرع من *InfoPro Technologies Service* و *CompuServe Information Dow* و *I.P. Sharp Associates* و *Data-Star* و *Dialog Information Services* و *Jones News/Retrieval* و *Mead Data Central* والمكتبة القومية للطب *National Library of Medicine* و *STN International* و *Questel* و *OCLC* و *VU/TEXT*

ويتيح بعض المتعهدين إمكانية الإفادة من أنواع شتى من مرصد البيانات (كما هو الحال مثلاً بالنسبة لـ *Dialog* ومؤسسة الاسترجاع الوراق *BRS*) بينما يتخصص آخرون ، سواء كان هذا التخصص في المجال الموضوعي (كما المكتبة القومية للطب *NLM* التي تتيح فرصة التعامل مع مرصد البيانات المتصلة بالطب) أو في نوعية مرصد البيانات التي تتيح فرصة الإفادة منها (فمؤسسة *Mead* تتعامل أساساً مع مرصد بيانات النصوص الكاملة ، في حين يتخصص مركز الحاسب الآلي للمكتبات على الخط المباشر *OCLC* في البيانات الوراقية) أو في جمهور المستفيدين المستهدف (حيث تركز مؤسسة *CompuServe* على المستفيدين الذين

لديهم حاسبات آلية شخصية بمنازلهم) . إلا أنه يحدث في كثير من الأحيان أن يتيح المتعهد إمكانية التعامل مع مرصد بيانات واحد فقط ، وهو المرصد الذي ينتجه .

ويقوم المتعهدون بتجهيز مراصد البيانات ، حيث يكشفونها بمختلف الطرق . وتوضح الأمثلة الواردة في الأشكال من ٩ إلى ١٢ مداخل مختارة من مختلف أنواع مراصد البيانات المتاحة عن طريق ديالوج . وفضلاً عن إتاحة فرصة وصول المستخدمين إلى البيانات في كل تسجيلية من التسجيلات عن طريق الحقول كالعنوان والمؤلف (الشكلان رقما ٨٩ و ١٠) واسم المترجم له (الشكل رقم ٩ B) والرقم الشفري للمنتج (الشكل رقم ١١) غالباً ما يقوم المتعهدون أيضاً بتجهيز محتوى كل حقل ، حيث يمكن على سبيل المثال اقتباس الكلمات المفردة وإضافتها إلى كشافات مراصد البيانات الوراقية ومراصد النصوص الكاملة .

ولما كانت كل من مراصد البيانات وخدمات البحث والاسترجاع قد نمت وتطورت مستقلة عن بعضها البعض إلى حد ما ، فإنه لا عجب أن يكون هناك تنوع كبير في سبل تطوير منتجات المعلومات وتجهيزها والتعامل معها . ويلخص وليامز (Williams 1980B) المشكلات المترتبة على غياب التقييس وتوحيد الموصفات في صناعة مراصد البيانات . فالوصول إلى المعلومات يفتقر إلى الموصفات الموحدة ، وبشكل واضح ، على النحو التالي :

١ - لمنتجي مراصد البيانات صيغهم الخاصة لتحديد بنية التسجيلات ومحتواها . فمن الممكن للتسجيلات في مختلف مراصد البيانات أن تختلف في طريقة الإشارة إلى المفهوم نفسه أو الموضوع أو اسم الشخص نفسه (ويشار إلى ذلك بغياب تقنين اللغة أو التقنين الاستنادي) . هذا بالإضافة إلى أنه من الممكن للتسجيلات في مرصد البيانات نفسه أن تختلف فيما بينها في طريقة التعبير عن اسم الشخص .

٢ - يتبع متعهدو مراصد البيانات طرقاً مختلفة في تجهيز المراصد .

٣ - تختلف مجموعات التعليمات (أسماء الأوامر ونظمها) الخاصة بالبحث في مراصد البيانات من متعهد إلى آخر .

ويوضح الشكلان رقما (١٣ ، ١٤) هذه النقاط ؛ فيشتمل الشكل رقم (١٣) على نماذج تسجيلات خاصة بوثيقتين مختلفتين ، حيث يبين العمود رقم (١) ثلاثة

A. Psycinfo

Document I

00700633 78-37258

Light therapy ■ seasonal affective disorder: A review of efficacy.

Terman, Michael; Terman, Juan S.; Quelin, Frederic M.; McGrath, Patrick J. et al

Columbia U, New York State Psychiatric Inst, US

Neuropsychopharmacology

1989 Mar Vol 2(1) 1-22

Coden: NEROEW ISSN: 0893-133X

Journal Announcement: 7811

Language: English

Document Type: JOURNAL ARTICLE

Composite Age: ADULT

Major Descriptors: 'ILLUMINATION; 'TREATMENT; 'SEASONAL VARIATIONS;

'AFFECTIVE DISTURBANCES

Minor Descriptors: ADULTHOOD

Descriptor Codes: 24420; 54180; 46030; 01260; 01150

Identifiers: light therapy efficacy, patients with seasonal affective disorder

Section Headings: 3300 -TREATMENT AND PREVENTION

II. MEDLINE

07118743 90028743

Light therapy for seasonal affective disorder: A review of efficacy.

Terman M; Terman JS; Quelin FM; McGrath PJ; Stewart JW; Ralston B

Department of Psychiatry, Columbia University, New York, NY.

Neuropsychopharmacology Mar 1989, 2 (1) pt 22, ISSN 0893-133X

Journal Code: ADQ

Contract/Grant No.: KO2 MH00461; RO1 MH42831; MHCRIC 30906

Language: ENGLISH

Document type: CLINICAL TRIAL; JOURNAL ARTICLE; MULTICENTER STUDY

Document II

00588589 74-00475

A circadian pacemaker for visual sensitivity?

Terman, Michael; Terman, Juan

Columbia U, New York State Psychiatric Inst

Annals of the New York Academy of Sciences

1985 Vol 453 147-161

Coden: ANYAAS ISSN: 00778623

Journal Announcement: 7401

Language: English

Document Type: JOURNAL ARTICLE

Major Descriptors: 'VISUAL PERCEPTION; 'BIOLOGICAL RHYTHMS;

'PROFESSIONAL MEETINGS AND SYMPOSIA

Minor Descriptors: RATS

Descriptor Codes: 35980; 06980; 40740; 42930

Identifiers: interactions between visual sensory & circadian functions, male rats, conference presentation

Section Headings: 2620 -NEUROLOGY & ELECTROPHYSIOLOGY

JOURNAL ANNOUNCEMENT: 8002

Subtitle: INDEX MEDICUS

Tags: Female; Human; Male; Support, U.S. Gov't, P.H.S.

Description: *Affective Disorders--Therapy--TH; Phototherapy; Circadian Rhythm; Multicenter Studies; Seasons

C. EMBASE

7560671 EMBASE No: 89126194

Light therapy for Seasonal Affective Disorder: A review of efficacy

Termen M.; Termen J.S.; Quilkin F.M.; McGrath P.J.; Stewart J.W.; Raderdy B.

**Department of Psychiatry, Columbia University, Columbia, NY USA
NEUROPSYCHOPHARMACOLOGY (USA) , 1989, 2/1 (1-22) CODEN: NERDCE ISSN:
0893-133X**

LANGUAGES: English

SUBFILES: 002

EMTAGS:

Psychological aspects 0136; Therapy 0160; Short survey 0002; Human 0888

MEDICAL DESCRIPTORS:

***affective neurasthenia--therapy--th; phototherapy
circadian rhythm; seasonal variation; statistics**

طرق مختلفة للتعبير عن وثيقة واحدة ، من جانب ثلاث مؤسسات مختلفة (الجمعية الأمريكية لعلم النفس American Psychological Association والمكتبة القومية للطب National Library of Medicine والمقتطفات الطبية Excerpta Medica) وبين العمود رقم (٢) تسجيلية لوثيقة أخرى في مرصد بيانات المعلومات النفسية PsycInfo . وتبين عناصر البيانات المقتطفة من هذه التسجيلات والواردة في الشكل رقم (١١٤) بوضوح غياب المواصفات الموحدة في التعبير عن كل من الموضوع واسم المؤلف ؛ فقد تم التعبير عن موضوع « الاضطرابات العاطفية الموسمية » Seasonal Affective Disorders « اللازم للدلالة على المفهوم الوارد في الوثيقة الأولية نفسها بثلاثة توافيق مختلفة للمصطلحات الكشفية . كما أن أحد المؤلفين (Jiuan S. Terman) لم يكتب اسمه بشكل مطرد في تسجيلات المنتجين الثلاثة ، وكذلك من جانب المنتج نفسه عند وصف وثيقتين مختلفتين للشخص نفسه . (إلا أن منتجي مرصد البيانات غالباً ما يقبلون شكل اسم المؤلف كما يرد في المصدر الذي يقومون بتكشيفه أو استخلاصه ، ومن ثم فإن جانباً من عدم الإطراد يرجع إلى الناشر الأولي ، وربما إلى المؤلف نفسه في بعض الأحيان) .

عناصر حقل الموضوع والمؤلف ١

| الوثيقة ٢ | الوثيقة ١ | | مرصد البيانات |
|--|------------------|---|-------------------|
| | المؤلف | الموضوع | |
| Terman, Jiuan S. | Terman, Jiuan S. | Affective disturbances Seasonal variations | PSYCINFO (a) |
| Terman JS | Terman JS | Affective Disorders Seasons | MEDLINE (b) |
| Terman J.S. | Terman J.S. | Affective neurosis Seasonal variation | EMBASE (c) |
| التجهيز من جانب المتعهدين ومجموعات التعليمات ب | | | |
| صيغة البحث | التجهيز | المتعهد | |
| affective disorders | اختر | تدخل الوصفات في الكشف كعبارات | DIALOG |
| affective-disorders | ابحث عن | تربط الوصفات بواصلات وتدخل الكشافات | BRS |

الشكل رقم (١٤) : نماذج غياب المواصفات الموحدة في نظم المعلومات .

ويوضح الشكل رقم (١٤ ب) غياب الالتزام بالمواصفات الموحدة من جانب المتعهدين . وأبرز مظاهر الاختلاف فيما بينهم أسماء الأوامر ، إلا أن هناك أيضاً اختلافات في التجهيز جديرة بالاعتبار ، كوجود الواصلات بين الواصفات في BRS وغيابها في DIALOG .

وعلى المستفيدين مراعاة عدم الالتزام بالمواصفات الموحدة هذا عند التفاعل مع نظم المعلومات ، إلا أنه يضيف ولا شك إلى تعقيدات عملية البحث . ولم يكن من الصعب التغلب على هذه المشكلة عندما كان وسطاء البحث (كالمكتبيين) هم الذين يفيدون بكثافة من عدد محدود من النظم ومراسد البيانات ، حيث كان بإمكانهم تعلم البحث في كل هذه النظم والمراسد بشكل فعال . إلا أن تزايد أعداد المتعهدين والمنتجين يجعل من الصعب إن لم يكن من المستحيل على الوسيط الإلمام بكل مراسد البيانات والمتعهدين ، حتى وإن كان ذلك في مجال موضوعي واحد . هذا بالإضافة إلى أن المستفيدين النهائيين يشكلون الآن قطاعاً متنامياً في السوق ، كما أنهم غالباً ما يفيدون من النظم على فترات متباعدة مما لا يتيح لهم فرصة تذكر لاكتساب المزيد من الخبرات الفنية .

وهناك في الأساس سبيلان للتعامل مع مشكلة التنوع والاختلاف :

١ - القضاء على الاختلاف في المنبع .

٢ - جعل الاختلاف ظاهراً (شفافاً) للمستفيد .

ويمكن للقضاء على الاختلاف في المنبع أن ينطوي على اتفاق جميع المنتجين على توحيد مواصفات محتوى مختلف فئات مراسد البيانات وبنيتها (كالمراسد الوراقية مثلاً) . وهذا في الأساس ما حدث فعلاً بالنسبة لمحتوى بيانات فهارس المكتبات ، من خلال الالتزام بتقنيات الفهرسة وبنية تسجيلات مارك ، إلا أن تسجيلات مارك هذه عادة ما يتم تجهيزها وتهيتها للبحث بطرق متعددة من جانب مختلف برمجيات الفهرسة على الخط المباشر . ولم تحقق مطالبة المتعهدين التجاريين بتوحيد مواصفات عمليات التجهيز ومجموعات تعليمات البحث الخاصة بهم ، أيضاً سوى نجاح محدود ، على الرغم من تكليف إحدى اللجان الفرعية بالمؤسسة القومية للمواصفات الموحدة في المعلومات (National Information Standards Organization (NISO) بوضع لغة أوامر موحدة ، لتوحيد مواصفات مجموعات التعليمات اللازمة للبحث في مختلف نظم الاسترجاع على الخط المباشر (Morrison, 1986) .

وينطوي جعل مشكلة الاختلاف ظاهرة أو شفافة على تصميم واجهات تعامل موحدة ، بحيث لا يحتاج المستخدم إلى تذكر أوجه الاختلاف والتعامل معها في إجراءات الربط بالخط المباشر logon ، ونظم الأوامر ، والحقول القابلة للبحث ، والتعبير عن البيانات كالاسماء والموضوعات . وأمامنا هنا عدة بدائل (Tenopir, 1986) :

١ - نظم المستخدم النهائي : وهذه يتم إعدادها وتطويرها بواسطة المتعهدين ، وتقدم التعليمات والأوامر اللازمة للبحث في مجموعات مراصد البيانات المتاحة على نظمهم المضيفة ، إما بقوائم الاختيار Menu - driven وإما بصيغ مبسطة . ومن أمثلة هذه النظم (BRS / Afterdark (BRS و Knowledge Index (DIALOG .

٢ - برمجيات الواجهة المباشرة Front - end : وهذه عبارة عن حزم برمجيات الحاسبات الآلية متناهية الصغر ، تساعد المستخدمين في صياغة استراتيجيات البحث ، خارج الخط المباشر عادة ، ثم تتصل تلقائياً بالنظام المضيف لتحميل الاستراتيجيات وإجراء عمليات البحث . ومن أمثلة هذه البرمجيات WILSEARCH و PRO-SEARCH .

٣ - بوابات المرور Gateways : وهذه تربط المستخدم عن طريق جهاز عملاق مضيف آخر (خدمة بوابة المرور) بمتعدد واحد أو أكثر . وتقدم كثير من بوابات المرور أيضاً بعض الخصائص الآلية المعاونة في البحث التي تساعد في صياغة استراتيجيات البحث واختيار مرصد بيانات المتعدد . ومن أشهر نماذج مثل هذه البوابات EASYNET .

وهناك أيضاً العديد من نظم برمجيات الواجهات المباشرة الخاصة بفهارس المكتبات ، حيث المستخدمون النهائيون هم أكثر الفئات تعاملأ مع هذه النظم .

والقدرة على إجراء البحث بيسر إحدى المزايا المهمة لهذه النظم ، ولكنها تنقلب إلى عيب جوهري وقضية أساسية في تطوراتها الراهنة . فنظراً لأنه من السهل نسبياً صياغة استراتيجية بحث ، فإن المستخدمين يمكن أن يعتقدوا أن عمليات البحث التي يجرونها أنجح مما هي في الواقع (Smith, 1986) . ويظل الموقف كذلك نظراً لأن هذه النظم المتعاطفة مع المستخدم تستمر في المعالجة الآلية للجوانب اليسيرة نسبياً فقط في اختيار مراصد البيانات وفي تطوير استراتيجيات البحث الفردية . أما المشكلات الأصعب من ذلك ، كالتعديل الآلي للمصطلحات في استراتيجية البحث بناء على النتائج

السابقة (Ide and Salton 1971) ، وتحويل المصطلحات من تلك المستعملة في أحد المراسد إلى مرصد آخر (Niehoff and Mack, 1985) ما زالت قيد الاستكشاف في جهود البحث العلمي في نظم استرجاع المعلومات .

نظم الاسترجاع على الخط المباشر والفهارس على الخط المباشر :

من بين أهداف هذا الكتاب توضيح وتوسعة التعريف المستقر لنظام اختزان المعلومات واسترجاعها . وقد تناول مجال المكتبات وعلم المعلومات فهارس المكتبات ، تاريخياً ، باعتبارها بيئة استرجاع مختلفة تماماً ، ولا ترتبط بصناعة مراسد البيانات . ولم يكن ذلك بالأمر المناسب ، لأنه أسفر عن الفرقة والعزلة في الإنتاج الفكري ، وهو موقف ما زال سائداً ، وإلى حد ما ، حتى يومنا هذا . وربما كان من الممكن تفهم ذلك إذا سلمنا بتتابع الأحداث الذي يصوره بوتير (Potter (1989 :

« لا تكفل معظم الفهارس الآن الوصول إلا إلى الكتب ومجموعات المسلسلات ، حيث يتعين على القراء المهتمين بالأنواع الأخرى للوثائق الرجوع إلى الكشافات المنفصلة عن الفهارس ، كالكشافات المطبوعة أو مراسد البيانات المسجلة على الأسطوانات الضوئية المكتنزة ، أو خدمات الاسترجاع على الخط المباشر التجارية » (p. 99) .

وربما يكون الإنتاج الفكري قد عبر إذن عن انفصال فهارس المكتبات عن غيرها من منتجات المعلومات ، إلا أن هناك من المؤشرات ما يدل على أن الاختلافات الواردة في تتابع الأحداث بين « الفهرس » و « الكشاف » وما يحيط بهما من بنية أساسية تقنية ، تتلاشى تدريجياً . وهناك اتجاهات عدة موثقة في الإنتاج الفكري ، ما زالت لها السيادة وبقوة ، وتسهم في إلقاء الضوء على احتمالات تطور هذا الموقف .

لقد كان هناك ، تاريخياً ، اختلاف بين المتعهدين مثل ديالوج الذين كانوا يقدمون مراسد البيانات في شكل قابل للبحث ، والمرافق الوراقية مثل مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC الذي كان يقدم مراسد بيانات التسجيلات اللازمة لإنتاج مداخل الفهارس . وتضطلع المرافق الوراقية الآن بمزيد من الأدوار المرتبطة بمتعهدي الخط المباشر التجاريين وبوابات المرور . فقد كانت وظائف البحث الخاصة بمركز الحاسب الآلي للمكتبات على الخط المباشر OCLC مثلاً تقتصر على عرض تسجيلات مارك للحصول على تسجيلات للفهارس ، إلا أنه من الممكن الآن البحث في

مرصد بيانات تسجيلات مارك عن طريق لغة استرجاع باللغة القوة (إبك EPIC) بكل إمكانات الاسترجاع التي يتميز بها المتعهدون الآخرون (Whitcomb, 1990) . هذا بالإضافة إلى إنتاج مركز الحاسب الآلي للمكتبات على الخط المباشر لبعض الاسطوانات الضوئية المكتنزة المشتملة على تسجيلات مارك ومرصد بيانات إريك ERIC ، منذ منتصف ثمانينيات القرن العشرين ، كما شرع الآن في تحميل بعض مراصد البيانات كإريك و *GPO Monthly Catalog* على حاسبه الآلي العملاق . وفي عام ١٩٩١م أدخل المركز ناتجاً جديداً ، وهو *First Search TM* المصمم أساساً لصالح المستفيدين النهائيين . ويتيح هذا الناتج إمكانية التعامل مع *OCLC Online Union Catalog* فضلاً عن العديد من مراصد البيانات الوراقية الأخرى والادوات المرجعية التي تنتجها الأجهزة الحكومية والتجارية .

وكانت مجالات الإفادة من تسجيلات فهرسة مارك والتسجيلات التي يعدها المنتجون التجاريون واضحة بما فيه الكفاية ، إلا أن ذلك قد بدأ أيضاً يتغير . فهناك عدد كبير من المتعهدين التجاريين الذين يتيحون إمكانية الاتصال بمراصد بيانات مارك (بما في ذلك تلك التي تقدمها مكتبة الكونجرس والمكتبة البريطانية) وهناك أيضاً تزايد في تحميل فهرس الخط المباشر للكشافات الإلكترونية لمقالات الدوريات والمجموعات الخاصة ، والنصوص الكاملة لبعض المواد كالموسوعات ، وتهيئة كل ذلك للبحث والاسترجاع (Potter, 1989) . هذا بالإضافة إلى أنه قد أصبح من الممكن الآن استرجاع المعلومات من أنواع شتى من مصادر المعلومات المتباينة ، والتي يمكن الوصول إليها الآن عن طريق انترنت *Internet* وهي شبكة عالمية لشبكات الحاسبات الآلية (Krol, 1992) فمن أي نقطة ارتكازية أو « مضيف » على انترنت يمكن للمستفيد الوصول إلى أنواع شتى من منتجات المعلومات المحملة على النظم المضيفة الأخرى، بما في ذلك مراصد البيانات الوراقية التقليدية ، وفهارس المكتبات على الخط المباشر ، والنصوص الكاملة للدوريات والكتب ، والصور وغير ذلك من مصادر المعلومات . كذلك تتوافر برمجيات البحث والاسترجاع مثل جوفر *Gopher* و وايز *WAIS* (*Wide Area Information Server*) و *WWW* (*World Wide Web*) في كثير من الحالات ، مما يتيح إمكانية الوصول إلى هذه المعلومات عبر واجهة تعامل موحدة أياً كانت نوعية مرصد البيانات أو بنيته .

وربما كانت أهم النقاط على الإطلاق هي أنه من الممكن توجيه الأسئلة نفسها التي أثرت حول نظم الاسترجاع التجارية والتجريبية ، والتي تتيح إمكانية الوصول إلى مقالات الدوريات أساساً ، توجيه هذه الأسئلة بالنسبة لفهارس المكتبات . وتتعلق هذه الأسئلة بقضية كفاءة الاسترجاع والعوامل المؤثرة فيها ، من لغة التكشيف ، وسياسة التكشيف ، وعدد التسجيلات التي يضمها مرصد البيانات ... إلى آخر ذلك من عوامل .

الفصل الثالث

الاسترجاع الموضوعي ؛ مشكلاته ومعايير الأداء

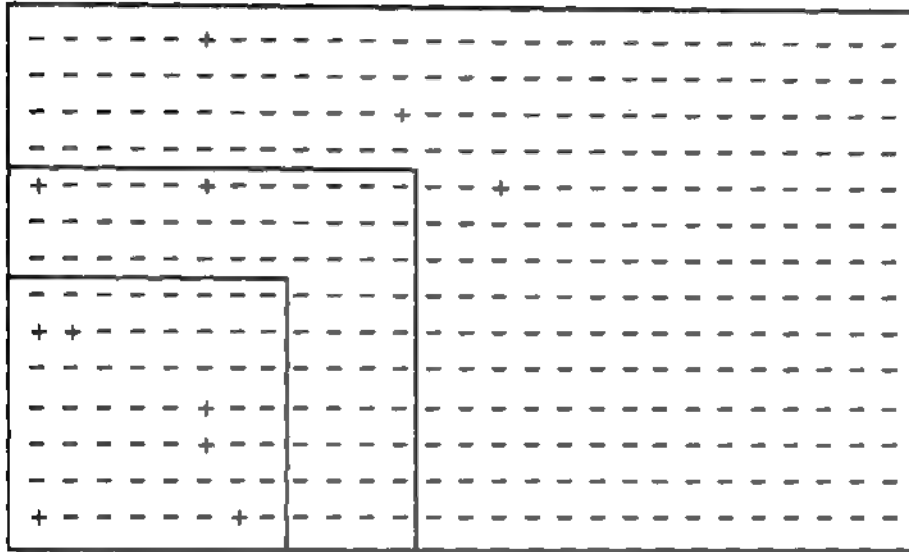
عادة ما يرغب المستفيد الذي يبحث في مرصد البيانات عن الوثائق التي تهتم بموضوع معين ، في العثور على بعض أو كل الوثائق التي تتناول الموضوع وتجنب استرجاع ماعداها . ويصدق ذلك بالنسبة لمرصد بيانات الكتب في مكتبة أو في مجموعة من المكتبات ، بالإضافة إلى مرصد البيانات التي تكفل الوصول إلى مقالات الدوريات وغيرها من أوعية المعلومات من شتى أنواع المصادر .

ومشكلة الاسترجاع الموضوعي أكثر تعقداً مما يمكن أن تبدو لأول وهلة . ونحاول تصويرها بيانياً في الشكل رقم (١٥) حيث يمثل المستطيل بأكمله مرصد البيانات والمواد التي يشتمل عليها . والمواد المشار إليها بعلامة « + » هي تلك التي يمكن للمستفيد الافتراضي أن يجدها مفيدة لتلبية الحاجة إلى المعلومات . أما المواد التي قد لا يراها المستفيد مفيدة فيشار إليها بعلامة « - » . وبالنسبة لأي موقف من مواقف الحاجة إلى المعلومات ، فإن عدد المواد المشار إليها بعلامة « - » يفوق بكثير عدد تلك المشار إليها بعلامة « + » . والواقع أن هذا الشكل لو نفذ وفقاً لمقياس الرسم لجاءت المواد الإحدى عشرة المفيدة مصحوبة بما يشغل مساحة جدار كامل من المواد غير المفيدة . والمشكلة هنا هي استرجاع أكبر عدد ممكن من الوثائق المفيدة وأقل عدد ممكن من الوثائق غير المفيدة .

مقاييس كفاءة الاسترجاع :

يمثل أصغر المستطيلين الداخليين في الشكل رقم (١٥) نتائج بحث تم في مرصد البيانات . وقد استرجع هذا البحث « ٥٧ » وثيقة ، كان من بينها ست وثائق مفيدة

و« ٥١ » وثيقة غير مفيدة . وتسمى نسبة الوثائق المفيدة إلى إجمالي الوثائق المسترجعة (٥٧ / ٦ ، أي حوالي ٠,٩ تقريباً) عادة نسبة التحقيق . أما النسبة التي تستعمل عادة للتعبير عن مدى العثور على جميع الوثائق المفيدة فهي نسبة الاستدعاء . ونسبة الاستدعاء في هذه الحالة ١١ / ٦ ، أي حوالي ٠,٥٤ تقريباً .



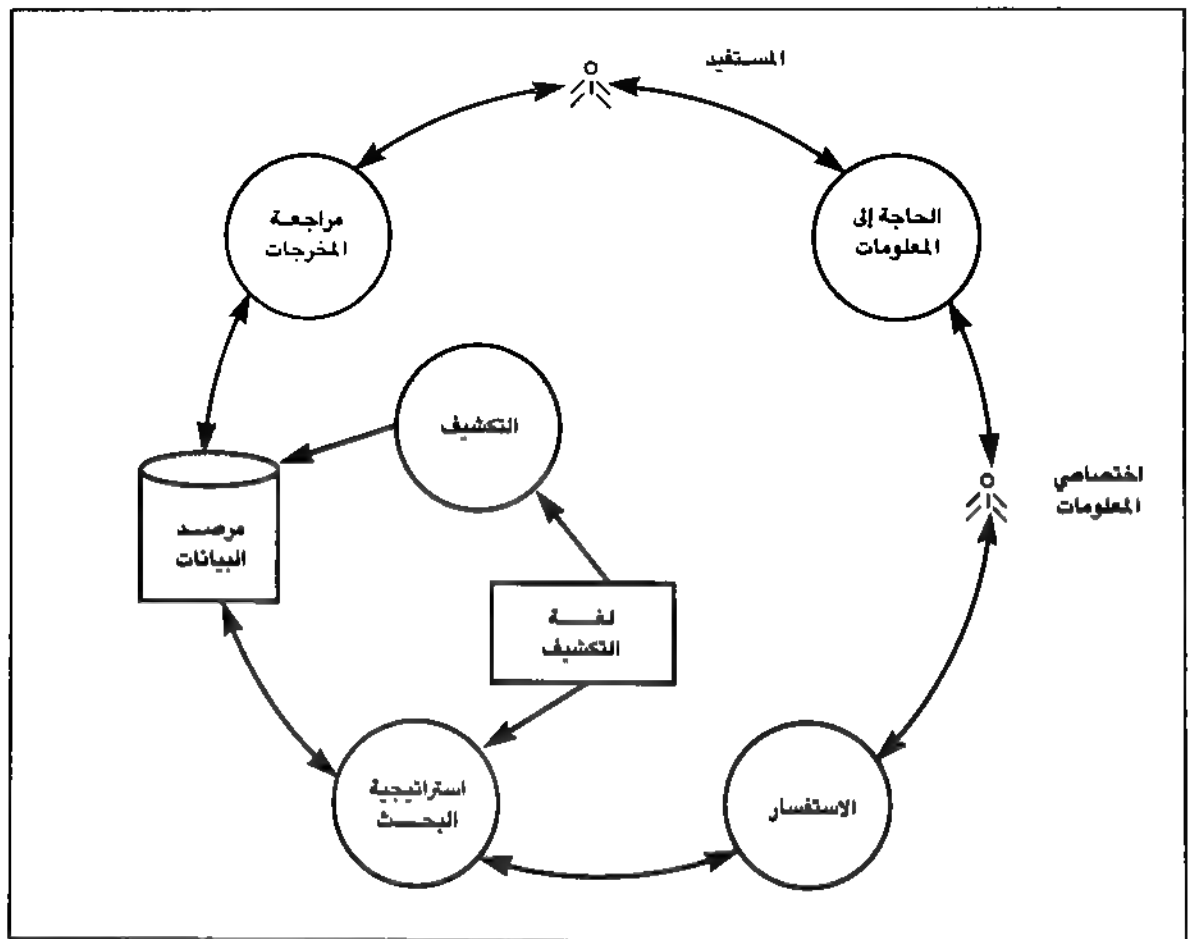
الشكل رقم (١٥) : مشكلة استرجاع الوثائق المناسبة من مرصد البيانات .

وللارتفاع بمستوى الاستدعاء في هذا الموقف ، فإننا قد نحتاج إلى البحث على نطاق أوسع . ويصور ذلك المستطيل الأكبر بين المستطيلين الداخليين ، حيث ارتفع الاستدعاء إلى ١١ / ٨ ، إلا أن التحقيق انخفض إلى ١١٢ / ٨ ، أي حوالي ٠,٠٧ تقريباً . ومن الخصائص المؤسسة لنظم استرجاع المعلومات أن الارتفاع بمستوى الاستدعاء عادة ما يؤدي إلى انخفاض التحقيق وأن الارتفاع بمستوى التحقيق عادة ما يكون على حساب الاستدعاء .

ويدل الشكل رقم (١٥) على وجود ظاهرة أخرى ! فربما كان من الممكن توسعة البحث بما فيه الكفاية للعثور على جميع الوثائق المفيدة (أي بلوغ نسبة الاستدعاء ١٠٠٪) إلا أن التحقيق يمكن أن يصل إلى حد لا يحتمل . أضف إلى ذلك أنه كلما تضخم حجم

مرصد البيانات تضاءلت فرص تحمل انخفاض نسبة التحقيق . وعلى الرغم من أن المستفيد قد يكون راجباً أو على استعداد للنظر في مستخلصات « ٥٧ » وثيقة للعثور على ست وثائق مفيدة ، فإنه قد يعزف تماماً عن النظر في « ٥٧٠ » مستخلصاً للعثور على ستين وثيقة مفيدة . ومن ثم فإنه يصبح من الصعب بمكان في مرصد البيانات بالغة الضخامة بلوغ مستوى مقبول للاستدعاء في مستوى تحقيق يمكن تحمله . وهذه الظاهرة يمكن ملاحظتها الآن بوضوح في فهارس المكتبات الضخمة المتاحة على الخط المباشر ، والتي يمكن للبحث فيها برأس موضوع أو بكلمة مفتاحية في العنوان ، أن يؤدي إلى استرجاع عدة مئات من التسجيلات .

وفي هذا الكتاب يستعمل المصطلح « استدعاء » للدلالة على القدرة على استرجاع الوثائق المفيدة ، ويستعمل « التحقيق » للدلالة على تجنب الوثائق غير المفيدة . وناقش هذين المقياسين وغيرهما من مقاييس كفاءة عمليات البحث في مرصد البيانات ، بشيء من التفصيل في الفصل التاسع .

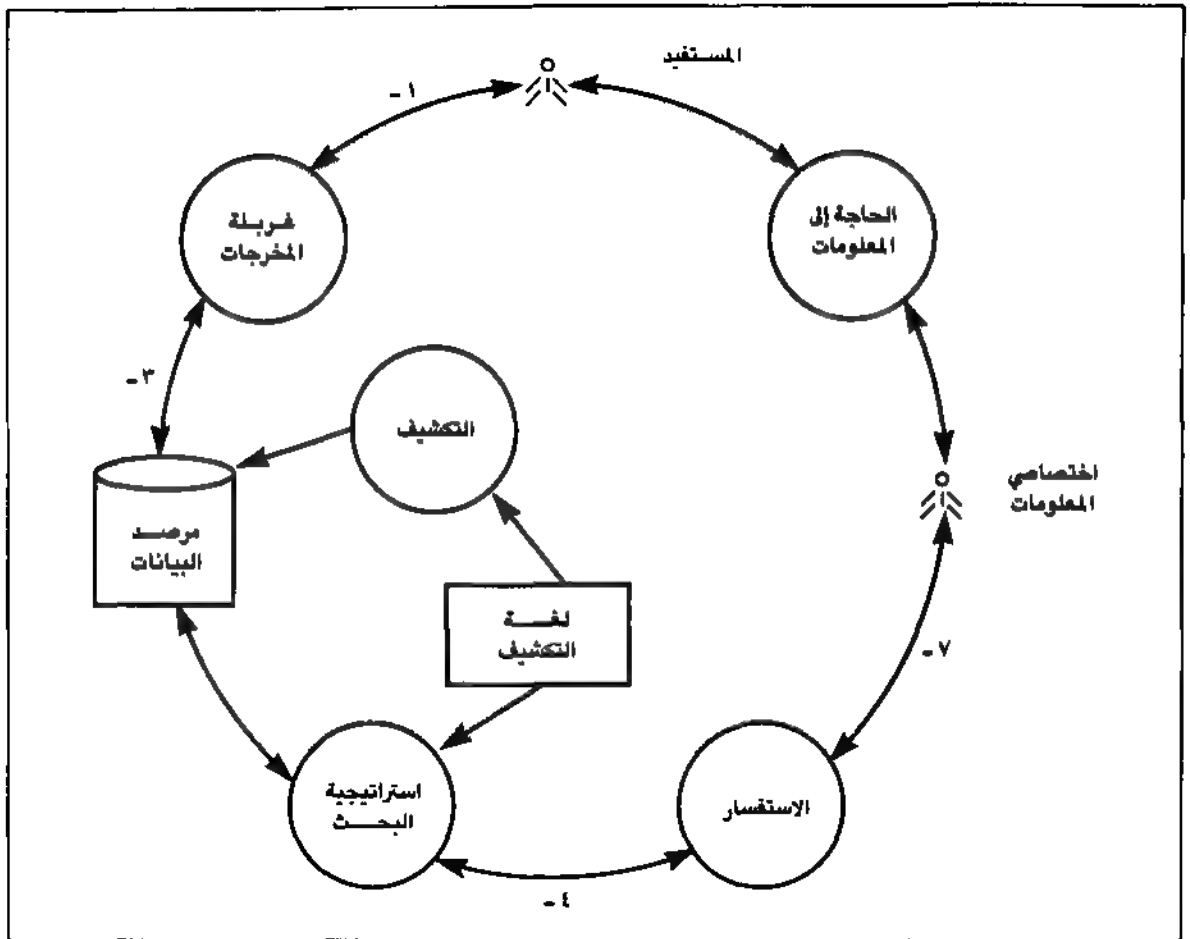


الشكل رقم (١٦) : العوامل المؤثرة في كفاءة البحث في مرصد البيانات .

العوامل المؤثرة في كفاءة نظم الاسترجاع :

هناك كثير من العوامل التي تؤثر في نجاح البحث في مرصد البيانات . ونحاول توضيح هذه العوامل على المستوى الكلي العام في الشكل رقم (١٦) . ويصور الرسم البياني موقفاً يلجأ فيه المستفيد من المكتبة أو مركز المعلومات إلى اختصاصي المعلومات طالباً البحث في مرصد بيانات معين . ويتم اتخاذ عدة خطوات قبل أن تسلم نتائج البحث إلى المستفيد ، ويمكن حدوث أخطاء في الاستدعاء أو التحقيق في كل خطوة من هذه الخطوات.

ففي البداية ينبغي أن يكون لدى اختصاصي المعلومات تصور واضح لما يبحث عنه المستفيد فعلاً . كما ينبغي نقل حاجة المستفيد إلى المعلومات إلى المسئول عن البحث في شكل استفسار . ولما تكون مثل هذه الاستفسارات تعبيراً دقيقاً عما وراءها من حاجة إلى المعلومات . ثم يحول الاستفسار بعد ذلك إلى استراتيجية بحث تلائم مرصد البيانات الذي سيتم البحث فيه . ويمكن لهذه الاستراتيجية أن يكتنفها بعض القصور ، وخاصة إذا لم يدرك اختصاصي المعلومات ما يريده المستفيد بوضوح ، كما يمكن لمسئول البحث أن يغفل عن بعض المصطلحات المهمة ، أو يستعمل بعض المصطلحات غير المناسبة فعلاً . هذا بالإضافة إلى أنه إذا ما استعمل مسئول البحث مصطلحات مستقاة من إحدى اللغات المقيدة فإن هذه المصطلحات قد لا تعبر عن الموضوع المطلوب بالشكل المناسب ، أو يحول بناء اللغة دون اهتداء مسئول البحث إلى جميع المصطلحات اللازمة للبحث الشامل . وعند مضاهاة استراتيجية البحث مقابل مرصد البيانات تبدأ جميع خصائص المرصد تؤدي دورها ؛ فمن الممكن أن تحدث أخطاء في الكشف ، حيث يحتمل ألا تكشف بعض الوثائق بالمصطلحات الكافية (وهذا أمر مرتبط بسياسة الكشف) كما يحتمل أن تكون اللغة المقيدة عاجزة عن التعبير عن الموضوع بالشكل الملائم (فقد لا تكون على سبيل المثال مخصصة بما فيه الكفاية) . ويمكن قبل تقديم النتائج للمستفيد أن يقوم مسئول البحث « بغربلتها » ولا يرسل إلا تلك التسجيلات التي تبدو في المقدمة من حيث الفائدة . ويتوقف نجاح عملية الغريple هذه على مدى إدراك مسئول البحث لحاجة المستفيد من المعلومات ، وعلى قدرته على التنبؤ اعتماداً على نتائج البحث (الإشارات الوراقية مثلاً أو الإشارات الوراقية مصحوبة بمستخلصات) بمدى صلاحية الوثائق .



الشكل رقم (١٧) : الفاقد في الاستدعاء في كل خطوة من خطوات سلسلة استرجاع المعلومات .

ويبين الشكل رقم (١٦) أسباب صعوبة إجراء عملية البحث التي تسفر عن استدعاء مرتفع وفي مستوى مقبول للتحقيق ، وخاصة في مرصد البيانات الضخمة . ومن الممكن إلقاء مزيد من الضوء على هذه الصعوبة بمثال (الشكل رقم ١٧) . لنفترض أن هناك في مرصد البيانات المزمع إجراء البحث فيه « ٣٠ » وثيقة يمكن أن يراها المستفيد مفيدة في تلبية حاجته من المعلومات . إلا أنه للأسف ، نظراً لقصور في استفسار المستفيد لن يتم استرجاع سبع من الثلاثين وثيقة . ويكتنف استراتيجية البحث شيء من القصور نظراً لأن مسئول البحث أغفل عدداً قليلاً من المصطلحات الملائمة ، ويؤدي ذلك إلى فقدان أربع وثائق أخرى مفيدة . وتحمل مظاهر القصور في تكشيف مرصد البيانات مسئولية المزيد من الإخفاق وضياح ثلاث وثائق أخرى . وأخيراً يعجز مسئول البحث عن موافاة المستفيد بإحدى الوثائق المسترجعة والتي كان من الممكن للمستفيد أن يحكم بصلاحياتها فعلاً ، وربما حدث ذلك لأن التسجيل الوراقية لا تدل على صلاحيتها

بوضوح . فمما لا شك فيه أن المستفيد لن يتلقى جميع الوثائق التي كان من الممكن أن تفيده ، وإنما يتلقى « ١٥ » وثيقة فقط منها (٣٠ — ٧ — ٤ — ٢ — ١) . وحينئذ تكون نسبة الاستدعاء لعملية البحث هذه ٠,٥ (٣٠ / ١٥) فقط نظراً لفقدان عدد من الوثائق في كل خطوة من خطوات السلسلة . ولنفس هذه العوامل ، يمكن أن يحدث إخفاق في التحقيق في كل خطوة (نتيجة مثلاً ، لافتقار لغة التكشيف إلى التخصيص أو نتيجة لمزيد من أخطاء التكشيف) ومن ثم فإن البحث يمكن ألا يتجاوز ٠,٥ في الاستدعاء و ٠,٤ في التحقيق مثلاً .

وكما يتضح من هذه المناقشة ، فإن العوامل الأساسية التي تؤثر في كفاءة البحث الموضوعي في مرصود البيانات ، عوامل فكرية وليست تقنية . وندقق هذه العوامل بشيء من التفصيل فيما يلي من فصول .

الصلاحية والاتصال بالموضوع :

غالباً ما يستعمل المصطلح « الاتصال بالموضوع relevance » للدلالة على ملائمة الوثيقة لمستفيد بعينه . إلا أنه ليس بالمصطلح الدقيق المحكم . ونحاول فيما بقي من هذا الفصل وضع حد فاصل له دلالتيه بين pertinence و relevance (*) ، المصطلحين اللذين يستعملان في الإنتاج الفكري لعلم المعلومات للدلالة على وجود علاقة بين الوثيقة والاستفسار عن المعلومات أو الحاجة إلى المعلومات أو الشخص الذي يطلب المعلومات أو يحتاج إليها . وهكذا يمكن القول بأن وثيقة ما تتصل باستفسار معين أو بحاجة إلى المعلومات ، أو بشخص يلتمس المعلومات عن موضوع معين ، أو تصلح لهذا الاستفسار أو هذه الحاجة أو هذا المستفيد . وللعلامة التي يعبر عنها هذان المصطلحان أهميتها البالغة في تقييم خدمات المعلومات . إلا أن كلا المصطلحين للأسف يستعملان في الإنتاج الفكري بشكل غير منضبط ، كما يثار كثير من الجدل حول معنى المصطلحين فعلاً ، وما إذا كان الاتصال بالموضوع relevance مناسباً أو وثيق الصلة بتقييم خدمات المعلومات . ويثير كل من ريس Rees وساراسفك (Saracevic 1966) ، على سبيل المثال التساؤلات التالية :

١ - هل الاتصال بالموضوع relevance معيار ملائم للقياس ، وهل هو المعيار الوحيد ؟

(*) كلاهما يقابله في العربية الاتصال بالموضوع ، ونميل لاعتبار pertinence مقابلًا للصلاحية . (المترجم) .

٢ - إذا كان الاتصال بالموضوع معياراً ملائماً ، فماذا يمثل ؟

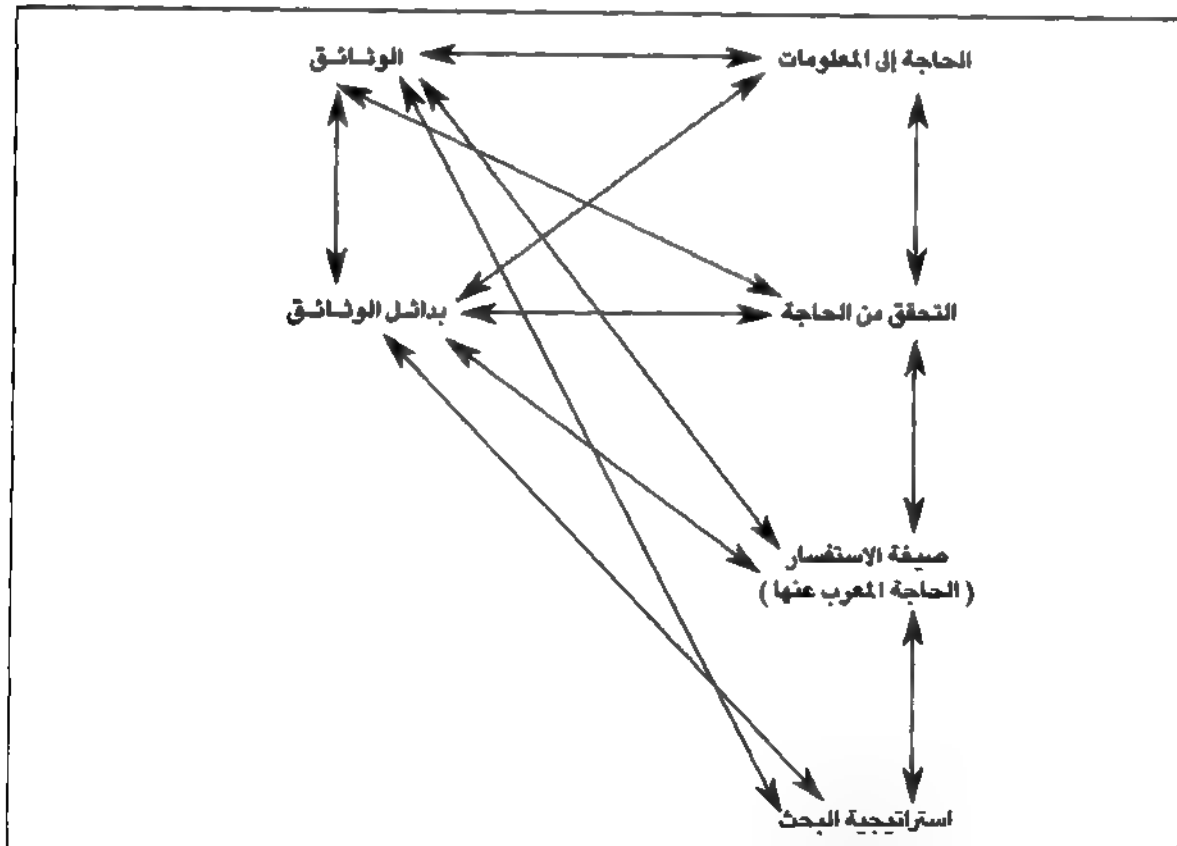
٣ - هل يمكن استخدام الاتصال بالموضوع في التطبيق العملي على الإطلاق ، هل هو قابل للقياس ؟ وإذا كان كذلك فعلاً ، ففي ظل أي الظروف (القيود أو المحاذير) يمكن استخدامه ؟ (P. 227) .

وينبغي النظر إلى مفهوم الاتصال بالموضوع في السياق العريض للشخص الذي يحتاج إلى المعلومات ويلجأ إلى أحد نظم الاسترجاع بحثاً عن هذه المعلومات . ومن المهم في هذا الصدد التمييز بين الحاجة إلى المعلومات ، والتحقق من هذه الحاجة ، والتعبير عنها . فنظم استرجاع المعلومات لا تستطيع الاستجابة لاحتياجات المستفيدين من المعلومات كاحتياجات ، وإنما لطرق التعبير عن هذه الاحتياجات فقط ، وعلى من يحتاج المعلومات أن يتعرف على حاجته ، وأن يكون لديه الحافز الكافي لاتخاذ تدابير تلبية هذه الحاجة . ولا يمكن التعبير عن الحاجة كطلب أو استفسار يقدم إلى مركز المعلومات ، إلا بعد أن يتعرف الفرد على الحاجة ويجد لديه الدافع الكافي . وعلى مدى دقة التحقق من طبيعة الحاجة إلى المعلومات ، ومدى دقة التعبير عن هذه الحاجة ، يتوقف وإلى حد بعيد نجاح مرفق المعلومات في إرضاء المستفيد . ويمكن لمرفق المعلومات أن يعمل تبعاً للاستفسار المقدم (الحاجة المعبر عنها) لا أكثر ، لأنه لا يستطيع ، ببساطة ، الاستجابة للاحتياجات التي لم يتم التحقق منها أو التعرف عليها ، أو الاحتياجات التي تم التعرف عليها ولم يتم التعبير أو الإعراب عنها . وكما نبين بمزيد من التفصيل في الفصل الرابع ، فإن من أهم التحديات التي يواجهها أي مرفق للمعلومات يعمل في ظل عمليات البحث المفوضة ، التأكد من أن الاحتياجات التي تم التعبير أو الإعراب عنها تطابق وبكل دقة الاحتياجات التي أمكن التعرف عليها أو التحقق منها . وليس من السهل في جميع الأحيان على من يحتاج إلى المعلومات التعبير عن تلك الحاجة بوضوح وبلا لبس ، لمن يضطلع بمسئولية البحث عن المعلومات .

دعنا نفترض إذن ، أن شخصاً ما يحتاج إلى المعلومات لجأ إلى مركز للمعلومات بحثاً عن هذه المعلومات ، ودعنا أيضاً نفترض أن هذا الشخص لا يحتاج إلى معطيات أو حقائق ، وإنما إلى الاطلاع على الوثائق التي تصف أو تناقش مجالاً موضوعياً بعينه ، وليكن مثلاً علاج المرض « هـ » بالعقار « سى » . ويحاول المستفيد التعبير عن هذه الحاجة للمسؤولين عن مرفق المعلومات في صيغة طلب أو استفسار . ويقوم أحد اختصاصيي

المعلومات بتحويل الاستفسار إلى استراتيجية بحث ، تتم مضاهاتها حينئذ مقابل مرصّد بيانات واحد أو أكثر . ونفترض أن البحث قد أجري في نظام إلكتروني وكانت مخرجاته قائمة مطبوعة ببدائل الوثائق ، كالإشارات الوراقية مثلاً ، الخاصة بالوثائق التي تضاهي استراتيجية البحث ، أي الوثائق التي كشفت بحيث تلبي المتطلبات المنطقية والمصطلحية للاستراتيجية . وأمامنا الآن مجموعة كاملة من العلاقات المحتملة التي ينبغي التعامل معها ، كما يصور الشكل رقم (١٨) .

ومما لا شك فيه أن بدائل الوثائق تضاهي استراتيجية البحث فعلاً ، وإلا ما كان لها أن تسترجع ، إلا أننا لسنا على يقين من مضاهاة الوثائق نفسها لاستراتيجية البحث؛ فمن الممكن أن يكون هناك خطأ في كشف بعض الوثائق . كما يمكن في حالات أخرى أن تكون المصطلحات التي أدت إلى استرجاع الوثيقة لا علاقة لها بالوثيقة في الأساس (ارتباط مزيف) أو تتصل بالوثيقة ولكن على نحو يختلف عما أراده مسئول البحث (علاقة غير صحيحة للمصطلح). ويمكن لبعض الوثائق المسترجعة أن تضاهي استراتيجية البحث ، ولكنها لا تضاهي صيغة السؤال أو الاستفسار (حاجة



الشكل رقم (١٨) : بعض العلاقات المهمة في تقييم نظم استرجاع المعلومات .

المستفيد التي أعرب عنها) . ويمكن لذلك أن يحدث ، على سبيل المثال ، إذا كانت استراتيجية البحث تشتمل على بعض المصطلحات غير الملائمة للسؤال ، أو إذا كانت عملية البحث قد أجريت على مستوى من التعميم أعلى مما كان محدداً في السؤال كذلك يمكن لبعض الوثائق المسترجعة أن تضاهي صيغة السؤال ولكنها لا تضاهي حاجة المستفيد كما تحقق منها . ويمكن لذلك أن يحدث عندما تعجز صيغة السؤال عن التعبير عن حاجة المستفيد التي أمكن التحقق منها ، بشكل كامل ودقيق . هذا ، ومن الممكن في النهاية ألا تضاهي بعض الوثائق كلاً من صيغة السؤال والحاجة التي أمكن التحقق منها ، إلا أنها تطابق تماماً ، وبمحض الاتفاق الحاجة الفعلية إلى المعلومات . ويمكن لذلك أن يعني أن النظام قد استرجع بعض الوثائق التي يراها المستفيد مفيدة في تلبية حاجته إلى المعلومات ، ولا تدخل هذه الوثائق في نطاق صيغة السؤال لأن المستفيد لم يكن يدري بوجود وثائق من هذا النوع . وعلى ذلك فإن الحاجة التي أمكن التحقق منها لم يكن من الممكن صياغتها بوضوح كاف ، يكفل تغطية هذه النوعية من الوثائق في صيغة السؤال أو الطلب .

وهكذا يتضح لنا وجود علاقات معقدة بين الحاجة إلى المعلومات ، والحاجة التي أمكن التحقق منها ، والحاجة المعرب عنها ، واستراتيجية البحث ، والوثائق ، وبدائل الوثائق . إلا أن الموقف برمته أكثر تعقداً نظراً لتغير بعض العلاقات بمرور الزمن . فكل من الاحتياجات التي أمكن التحقق منها ، والاحتياجات الفعلية يمكن أن تتغير . وعلى ذلك فإنه يمكن لوثيقة ما أن تضاهي حاجة المستفيد التي أمكن التحقق منها يوماً ولا تضاهيها بعد ذلك ، لأن إدراك المستفيد لحاجته قد تغير .

ولا ننسى أن كلمة يضاهي match المحايدة قد استعملت في هذه المناقشة للدلالة على مختلف العلاقات بين الطلبات أو الأسئلة ، والحاجة إلى المعلومات ، والوثائق ، وبدائل الوثائق . ومن الضروري الآن إمعان النظر في أنماط ما ينطوي عليه الموقف من مضاهاة ؛ من الذي يستطيع أن يقرر ما إذا كان نوع ما من المضاهاة قد حدث أم لا ، وما الأسماء المحتملة لهذه المضاهاة . وربما كان اختصاصي المعلومات الذي يقوم بإجراء عملية البحث ، أو أي اختصاصي معلومات آخر يرتبط بالنظام ، أقدر الجميع على أن يقرر ما إذا كانت بدائل الوثائق تضاهي استراتيجية البحث أم لا . و « المضاهاة »

مصطلح ملائم جداً بالفعل لهذه العلاقة : فبديل الوثيقة يضاهي إذا كان يشتمل على مصطلح معين ، أو مجموعة مؤتلفة من المصطلحات ، الواردة في استراتيجية البحث . ويمكن في الواقع ، لأي إنسان قادر على قراءة الاستراتيجية وإدراك منطقتها أن يقرر ما إذا كانت المضاهاة قد تمت أم لا .

وما إذا كانت وثيقة بعينها تضاهي استراتيجية البحث أم لا ، أحد الأمور التي يمكن أن يقررها على أحسن وجه من يقوم بوضع الاستراتيجية ، وإن كان من الممكن أن يتقرر أيضاً من جانب اختصاصي معلومات آخر . وهذا الموقف أكثر تعقداً من الآخر من ناحيتين :

١ - فنحن نهتم بما هو أكثر من مجرد مضاهاة مصطلح أو كلمة ، ونهتم هنا بما بين المصطلحات أو الكلمات من علاقات ، أي العلاقات النظمية والدلالية .

٢ - والمضاهاة التي نهتم بها ، إلى حد ما ، مضاهاة بين الوثيقة والاستراتيجية المقصودة . ولهذا فإن من يقوم بصياغة استراتيجية البحث هو الأقدر على أن يحكم ما إذا كانت المضاهاة المقصودة قد حدثت أم لا .

ويتضح هذا الموقف على خير وجه بمثال بسيط . فإمامنا الحالة التي يستعمل فيها مسئول البحث الاستراتيجية « **READING and EPILEPSY** » « القراءة والصرع » أي أن المسئول عن البحث يطلب الوثائق التي تشتمل على الكلمتين معاً ، أو الوثائق التي كشفت تحت كل من المصطلحين . ويتم استرجاع عدة تسجيلات تشتمل على المصطلحين أو الكلمتين ، وكلها تضاهي الاستراتيجية . فقد استرجع النظام ما طلب منه استرجاعه ، وتصرفت برامج البحث على أحسن وجه . إلا أن مسئول البحث عندما ينظر إلى مجموعة الوثائق المناظرة للتسجيلات المسترجعة يكتشف أنها تنقسم إلى فئتين :

١ - مجموعة كبيرة من الوثائق التي تناقش الصرع الناتج عن القراءة ، أي أحد أنواع الصرع الضوئي **Photic** .

٢ - مجموعة صغيرة من الوثائق التي تناقش القدرات القرائية للأطفال المصابين بالصرع . والفئة الأولى فقط هي ما قصد مسئول البحث استرجاعه ، أما الفئة الثانية فلم تكن مرغوبة وربما لم تكن متوقعة .

وفي هذه الحالة تضاهي جميع التسجيلات المسترجعة استراتيجية البحث ، إلا أن بعضها لا يضاهي الاستراتيجية المقصودة . ولا يكمن الخطأ في برامج البحث أو في تفسير مسئول البحث لاحتياجات المستفيد ، وإنما في إجراءات الكشف ولغة الكشف المستخدمة في النظام . فلو كانت إجراءات الكشف تحرص على تحديد ما بين المصطلحات من علاقات بدقة ، باستعمال مؤشرات الدور مثلاً (مثل القراءة (٤) والصراع (٢) ، حيث ٤ يمثل السبب و ٢ يمثل الأثر) ، أو كانت لغة الكشف أكثر تخصيصاً ، كأن يرد بالكنز مصطلح **READING EPILEPSY** صرع القراءة مثلاً ، لكان من الممكن تجنب هذه الوثائق غير المرغوبة . ومن الممكن افتراض موقف آخر ، تشتمل فيه تسجيلية وثيقة على المصطلح الكشفي « القراءة » والمصطلح الكشفي « الصرع » ، أي أنها تضاهي فعلاً استراتيجية البحث ، ولكنها تشير إلى وثيقة لا تتناول صرع القراءة وإنما شكلاً آخر للمرض . فقد أساء المكشف تفسير الوثيقة وأخطأ في كشفها . وفي هذه الحالة تضاهي التسجيلية ، كما هي ، الاستراتيجية ، إلا أن هذه التسجيلية غير صحيحة نتيجة لخطأ في الكشف .

وهذه العلاقات بين تسجيلات الوثائق واستراتيجيات البحث ، داخلية إلى حد ما ، بالنسبة لنظام الاسترجاع ، ولا تستدعي قرارات الحكم على مدى ملاءمة عمليات المضاهاة تدخل المستفيدين من النظام أو غيرهم من اختصاصيي المعلومات . ويحدث في بعض الأحيان أن يعجز غير المرتبطين مباشرة بالنظام عن إدراك السبب في حدوث خطأ معين أو حتى السبب في استرجاع وثيقة معينة . وربما كان من الأفضل ألا نستعمل الاتصال بالموضوع **relevance** أو الصلاحية **pertinence** بالنسبة لهذه العلاقات ، وإنما نشير ببساطة لتسجيلات أو بدائل الوثائق « التي تضاهي استراتيجية البحث » والوثائق « التي تضاهي الاستراتيجية المقصودة » .

الاتصال بالموضوع :

يهتم استرجاع المعلومات بما هو أكثر من المضاهاة بين الوثائق أو بدائلها واستراتيجيات البحث ؛ إنه يهتم بالعلاقات بين الوثائق وصيغ الطلب أو الاستفسار ، وبين الوثائق واحتياجات المستفيدين من المعلومات . فبعد أن يسترجع النظام مجموعة من تسجيلات الوثائق استجابة لطلب معين ، يمكن استرجاع الوثائق نفسها ، وتحديد

ما يضاهي منها الطلب وما لا يضاهيه . فمن المؤهل لاتخاذ هذا القرار؟ ربما يكون اختصاصيو المعلومات المرتبطون بالنظام ، أو المستفيدون ، أو اختصاصيون موضوعيون محايدون . ولا شك أنه يتعين على من يتخذ القرار أن يكون على دراية كافية بالموضوع حتى يكون قادراً على أن يؤكد أن هناك وثائق معينة تعد « استجابة صحيحة » للطلب وأخرى ليست كذلك .

ولا يمكن أن ننكر أن المصطلح « استجابة صحيحة » غامض . فالشخص الذي يتولى التقييم ، من المفترض أنه يحكم بأن وثيقة ما تعد استجابة صحيحة ، إذا اعتقد أن موضوع الوثيقة قريب بما فيه الكفاية من الموضوع المطلوب ، وأن النظام كان « محقاً » في استرجاعها . وهذه العلاقات للأسف ليست على قدر كبير من الانضباط ؛ فكيف يكون القريب قريباً بما فيه الكفاية ؟ إن مدى قرب الوثيقة من الطلب لا يمكن قياسه بدقة ؛ فالعلاقة ذاتية ومرنة وليست موضوعية وحاسمة ، ويمكن أن تختلف القرارات حول مدى الارتباط بين الوثيقة والطلب باختلاف المحكمين . كما أنه من الممكن أيضاً لنفس المحكم أن يتخذ قرارات مختلفة حول العلاقة بين نفس الوثيقة ونفس الطلب باختلاف الزمن . وإذا كان الأمر كذلك فعلاً ، فإنه يبدو من الأوفق الإشارة إلى هذه العلاقة باعتبارها الاتصال بالموضوع . وبعبارة أخرى ، فإنه إذا قرر أحد المحكمين أن وثيقة معينة قريبة بما فيه الكفاية في موضوعها إلى طلب معين ، وأن النظام قد أصاب في استرجاعها ، فإنه لا تجاوز على الإطلاق في القول بأنه حكم باتصال الوثيقة بموضوع الطلب .

وعلى الرغم من إمكان استعمال مصطلحات أخرى ، فإن المصطلح الاتصال بالموضوع **relevance** مناسب للدلالة على وجود علاقة بين الوثيقة وصيغة الطلب في نظر المحكم . وقد يكون من الخطأ التسليم بأن الاتصال بالموضوع يمثل علاقة محددة ثابتة ، وهو ليس كذلك فعلاً . وواقع الأمر أنه بدلاً من القول بأن الوثيقة تتصل بموضوع الطلب ربما كان من الأفضل القول بأن الوثيقة قد حكم بأنها تتصل بموضوع الطلب .

ونظراً لأن قرارات الاتصال بالموضوع ذاتية ويمكن أن تكون غير مطردة ، فإنه قد يكون من الخطأ الفادح استعمال مجموعة واحدة من قرارات الاتصال بالموضوع والتي

اتخذها فرد واحد ، كأساس لتقييم كفاءة نظام الاسترجاع . وإنما لكي نقيم كفاءة نظام الاسترجاع على أساس العلاقة بين الطلب والوثائق فقط ، فإنه ينبغي مشاركة مجموعة من المحكمين ليحاولوا التوصل إلى اتفاق حول أي الوثائق تتصل بأي الطلبات . ويمكن لعدة محكمين يعمل كل منهم مستقلاً عن الآخر في إتخاذ قرارات الاتصال بالموضوع . أن يتيحوا على الأقل إمكانية ترتيب الوثائق طبقاً وفقاً « للاتفاق حول الاتصال بالموضوع » . وحينئذ يمكن التعبير عن نتائج عملية بحث بعينها على النحو التالي :

١ - قرر المحكمون الخمسة بالإجماع أن ٣٥٪ من الوثائق المسترجعة تتصل بالموضوع .

٢ - قرر أربعة على الأقل من المحكمين الخمسة أن ٤٣٪ من الوثائق المسترجعة تتصل بالموضوع .

٣ - قرر ثلاثة على الأقل من المحكمين أن ٦٢٪ من الوثائق المسترجعة تتصل بالموضوع . وهكذا ...

الصلاحية :

ليس لقرارات الاتصال بالموضوع المعتمدة على العلاقة بين الوثائق وصيغ الطلبات سوى فائدة محدودة في تقييم مرافق المعلومات العاملة فعلاً ؛ فهي لا تقدم شيئاً عن مدى ما يتحقق من نجاح في تلبية احتياجات المستفيدين من المعلومات ، ومن المفترض طبعاً أن المرفق إنما وجد لتلبية مثل هذه الاحتياجات . ومن الممكن الحكم باتصال وثيقة ما بموضوع صيغة طلب معين ، من جانب جميع أعضاء مجموعة المحكمين ، إلا أنه من الممكن لمن قدم الطلب أن يقرر أن الوثيقة لا قيمة لها في تلبية الحاجة إلى المعلومات التي دفعت إلى تقديم الطلب إلى المرفق . وإذا حكم المستفيد بأن جميع الوثائق المسترجعة بعملية البحث لا قيمة لها ، فإنه سوف يرى عملية البحث غير موفقة ، أيأ ما كان عدد المحكمين الذين اتفقوا على أن الوثائق تتصل بموضوع صيغة الطلب .

ويمكن للاختلافات الجوهرية بين قرارات اتصال الوثائق بالطلب التي يتخذها مجموعة من المحكمين ، والقرارات الخاصة بما للوثائق من قيمة بالنسبة للحاجة إلى المعلومات والتي يتخذها المستفيدون ، أن تدل على أن صيغة الطلب (الحاجة المعرب عنها) والتي عمل النظام بناء عليها ، لم تكن تعبر عن الحاجة الفعلية إلى المعلومات تعبيراً

دقيقاً . ولا عجب في الواقع إن اختلفت مجموعات القرارات هذه اختلافاً كبيراً ، نظراً لأنه من الصعب في غالب الأحيان ، وحتى على المستفيد الواعي الحريص تعريف مرفق المعلومات بالاحتياجات الفعلية إلى المعلومات . فقد سجل لانكستر (1968 a) Lancaster على سبيل المثال ، في تقييمه لمرصّد بيانات المدلرز MEDLARS أن عملية بحث أجريت بناء على صيغة الطلب « السرطان في الأجنة أو الأطفال حديثي الولادة » قد استرجعت « ١١٦٧ » مقالة ، كان من الممكن الحكم باتصال الغالبية العظمى منها بموضوع صيغة الطلب من جانب أي مجموعة من الأطباء الممارسين . إلا أن المستفيد قدم طلباً سيئاً بلا مبرر للنظام ، طلباً أعم بكثير من الحاجة الفعلية إلى المعلومات . فقد كان المستفيد يهتم فعلاً بالعلاقة بين العيوب الخلقية Teratogenesis والأورام Oncogenesis على المستوى الخلوي Cellular ، وهو موضوع مغرق في التخصص ، وقرر أن وثيقة واحدة فقط من بين الـ « ١١٦٧ » وثيقة المسترجعة لها قيمة في تلبية حاجته إلى المعلومات .

وهناك في الإنتاج الفكري من يجادل بأنه لا ينبغي الحكم على النظام إلا بناء على العلاقة بين الطلب والوثيقة فقط ، وأنه لا ينبغي أن نتوقع منه إلا ما يطلب منه القيام به . فإذا طلب من النظام استرجاع الوثائق المتصلة بالسرطان في الأجنة أو الأطفال حديثي الولادة ، وفعل ذلك وفقاً لتقدير مجموعة من المحكمين ، فقد تصرف تصرفاً سليماً ، على الرغم مما قد يراه المستفيد في النتائج . وهذه نظرة قاصرة جداً ، نظرة تنطوي على تجاهل تام لمسئولية النظام عن التأكد قدر الإمكان من أن ما يتلقاه من طلبات تعبر بدقة عن حاجة المستفيدين من المعلومات . فالنظام الذي يقبل جميع الطلبات بمعناها الظاهري ، ويحكم على كفاءته بناء على قدرته على الاستجابة لها ، نظام محكوم عليه بالفشل المؤكد تقريباً .

ولتقييم مرفق معلومات « حقيقي » مرفق يستقبل مستفيدين حقيقيين ، يقدمون طلبات حقيقية بناء على احتياجات حقيقية إلى المعلومات ، فإنه لا مناص من أن يكون المعيار هو قدرة المرفق على تلبية احتياجات المستفيدين إلى المعلومات . ومما لا شك فيه أن المستفيد وحده هو القادر على أن يقرر ما إذا كانت وثيقة معينة تسهم في تلبية حاجته إلى المعلومات أم لا ، ذلك لأن المستفيد هو وحده الذي يعرف هذه الحاجة فعلاً . ومن الممكن النظر إلى هذه القرارات أيضاً باعتبارها قرارات اتصال بالموضوع . فهي

قرارات « اتصال بموضوع الحاجة إلى المعلومات » لا قرارات « اتصال بموضوع الطلب ». إلا أنه من الأوفق استعمال المصطلح صلاحية *pertinence* في هذه الحالة ؛ أي استعمال الاتصال بالموضوع *relevance* للدلالة على العلاقة بين الوثيقة والطلب بناء على القرار الذاتي الصادر عن فرد واحد أو أكثر ، واستعمال الصلاحية *pertinence* للدلالة على العلاقة بين الوثيقة والحاجة إلى المعلومات بناء على قرار شخص واحد وهو الشخص الذي يحتاج إلى المعلومات .

وبصراحة ، فإن أي المصطلحات تستعمل ليس بالأمر المهم ، وإنما المهم فعلاً أن نتحقق من الفرق بين علاقتين (علاقة الطلب بالوثيقة وعلاقة الحاجة إلى المعلومات بالوثيقة) وأن نتحقق ممن هو أهل لأن يتخذ القرار في كل من الحالتين ، وأننا نستعمل المصطلحات بشكل مطرد أياً كانت المصطلحات التي تروق لنا . ولتحقيق هذا الإطار في هذا الكتاب ، يستعمل المصطلح صلاحية *pertinence* للدلالة على العلاقة بين الوثيقة والحاجة إلى المعلومات ، والمصطلح الاتصال بالموضوع *relevance* للدلالة على العلاقة بين الوثيقة وصيغة الطلب . وهذا قرار تعسفي اتخذ لأجل الصلاحية في المقام الأول . (وربما يكون من المفضل فعلاً النظر في هذه العلاقات ، باعتبارها الاتصال بموضوع الطلب والاتصال بموضوع الحاجة إلى المعلومات ، مع استبعاد المصطلح *pertinence* كلية).

وقرارات الصلاحية كما عرفناها توأ ، أحكام قيمية يصدرها المستفيدون . ويعبر القرار عما للوثيقة من قيمة بالنسبة لشخص معين في وقت معين ، في الإسهام في تلبية حاجة معينة إلى المعلومات . ولا « تصدق » مجموعة الأحكام القيمية هذه إلا في مدى زمني معين ، وهو الوقت الذي يتم فيه اتخاذ القرار ، نظراً لأن الحاجة إلى المعلومات تتغير . وعلى ذلك ، فإنه يمكن للمستفيد أن يلجأ إلى أحد مرافق المعلومات ، ويقدم طلباً بناء على حاجته إلى المعلومات التي تحقق منها . وإذا افترضنا أن النظام قادر على الاستجابة الفورية تقريباً ، واسترجع « ٢٥ » وثيقة ، فإنه في اللحظة التي يرى فيها المستفيد الوثيقة الأولى ، يمكن لإدراكه للحاجة أن يتغير ، ويؤثر ذلك حتماً في جميع قرارات الصلاحية التي يصدرها . وقد تكون هذه الوثيقة الأولى هي الوثيقة التي يحتاجها المستفيد فعلاً على وجه التحديد ، فترضي حاجته إلى المعلومات تماماً . ومن

الممكن أن يحكم المستفيد على الوثيقة السادسة بأنها لا قيمة لها ، لأنها تكرر الأولى في الأساس . ولكن لو حدث أن اطلع المستفيد على الوثيقة السادسة قبل غيرها لكان من الممكن أن يحكم بصلاحياتها ، بينما الوثيقة الأولى لا قيمة لها إذا وردت بعد ذلك في الترتيب . أما إذا حدث أن تلقى المستفيد نتائج البحث بعد عدة أيام ، لا في نفس اليوم الذي قدم فيه الطلب ، فإن إدراك المستفيد للحاجة إلى المعلومات لا بد وأن يتأثر بكل ما تعلمه خلال الفترة الفاصلة . وليس من الضروري للأحكام القيمية التي تصدر على الوثائق الخمس والعشرين ، في اليوم العاشر أن تكون هي نفس الأحكام التي كان من الممكن أن تصدر في اليوم الأول ، كما يمكن لكل من أحكام اليوم العاشر واليوم الأول أن تختلف عن تلك التي كان من الممكن أن تصدر في اليوم الخامس (١) .

وقرارات الصلاحية إذن وقتية جداً ، وهي أكثر قابلية للتغير من قرارات الاتصال بالموضوع . وهي تتأثر بكل من عامل الزمن والتسلسل الذي يتم به اتخاذ القرارات . ولا يسهم ذلك في تيسير مهمة إدارة مرافق المعلومات ، وإنما يعد إحدى الحقائق التي يتعين على مصممي النظم ومديريها والمسؤولين عن تشغيلها والمهتمين بتقييمها ، إدراكها ، وأن يكونوا قادرين على التكيف معها .

يوضح الشكل رقم (١٩) مختلف العلاقات التي ناقشناها ، بالإضافة إلى بيان المؤهلين للحكم على ما إذا كانت هناك علاقة فعلاً أم لا ، والأسباب الكامنة وراء الحكم بوجود العلاقة أو غيابها . وقد وردت العلاقات الأربع مرتبة تصاعدياً وفقاً لمدى تعقدها . وفيما يتعلق بما لها من أهمية في تقييم النظام ، فإنه يمكن النظر إلى البيانات المعبرة عن العلاقات باعتبارها تراكمية . ومن الممكن توضيح هذه العبارة على النحو التالي :

١ - إذا كان من المعروف أي بدائل الوثائق التي استرجعها بحث معين تضاهي فعلاً استراتيجية البحث المستخدمة ، فإنه يمكن تتبع المشكلات في برامج البحث . وهذه في الواقع حالة لا وزن لها ولا أثر ، حيث لا يمكن توقع أخطاء من هذا النوع إلا في النظم فادحة الفشل . ولا يقدم هذا المستوى في التقييم شيئاً يذكر عن نوعية التكشيف أو لغة التكشيف أو استراتيجيات البحث .

(١) نشر أيزنبرج Eisenberg وبري Perry (1988) تقريراً عن دراسة حديثة تناولت أثر ترتيب تسجيلات الوثائق على أحكام الاتصال بالموضوع . إلا أن باركر وجونسون Parker and Johnson توصلا إلى نتائج تدل على أن ترتيب التسجيلات لا يؤثر في أحكام الاتصال بالموضوع إلا عندما يزيد عدد المواد التي يتم تقييمها عن حد معين . وهما يريان أن الوثائق التي تقدم قبل غيرها لا يمكن أن تكون فرصتها في الحكم باتصالها أفضل من تلك التي ترد فيما بعد ، إلا عندما يزيد مجموع الوثائق موضوع التقييم عن ١٥ وثيقة .

٢ - إذا كان من المعروف أي الوثائق المناظرة للبدائل المسترجعة عن طريق البحث، ترضي الاستراتيجية المقصودة للمسئول عن البحث ، فإنه يمكن تتبع بعض المشكلات في برامج البحث ، بالإضافة إلى بعض مشكلات الكشف ، أو لغة الكشف أو استراتيجية البحث ، كما يتضح من الشكل رقم (١٩). إلا أن هذا الموقف لا يقدم لنا شيئاً يذكر عن نوعية استراتيجية البحث باعتبارها تعبيراً عن صيغة الطلب أو بديلاً لها .

٣ - إذا كان من المعروف أي الوثائق المناظرة للتسجيلات المسترجعة بالبحث قد أقر اتصالها بموضوع صيغة الطلب ، من جانب مجموعة محكمين من الاختصاصيين الموضوعيين ، فإنه يمكن تتبع المشكلات في برامج البحث ، وفي الكشف ، ولغة الكشف ، واستراتيجيات البحث . كذلك يمكن أيضاً تتبع بعض الحالات التي لا تعبر فيها استراتيجية البحث ، بشكل مكتمل أو دقيق ، عن مضمون صيغة الطلب . إلا أننا مازلنا لا نعرف شيئاً عن مدى قدرة الوثائق المسترجعة على تلبية حاجة المستفيد إلى المعلومات ، ومن ثم فإنه لا يمكن التحقق من صيغ الطلب غير الملائمة أو المضللة ، نتيجة القصور في تفاعل المستفيد مع النظام .

٤ - إذا كان من المعروف أي الوثائق المسترجعة قد قرر المستفيد أنها تسهم فعلاً في تلبية حاجته إلى المعلومات ، أي الوثائق الصالحة ، وإذا كان من المعروف أيضاً أسباب الحكم بعدم صلاحية ، فإنه يمكن حينئذ التمييز بين اتصال نتائج البحث بموضوعه من جهة وصلاحية هذه النتائج من جهة أخرى . ومن الممكن في هذا الموقف ، وعن طريق أساليب التحليل المناسبة التحقق من جميع أنواع المشكلات التي يمكن أن تحدث في نظام الاسترجاع ، سواء في برامج البحث ، أو في الكشف ، أو في لغة الكشف ، أو في استراتيجية البحث ، أو في تفسير الطلب من جانب مسئول البحث ، أو في تفاعل المستفيد مع النظام . ومن الأمثلة على استعمال هذا النوع من التحليل التشخيصي تقييم مرصد بيانات المدلرز (Lancaster, 1968 a)

إلا أننا ينبغي ألا ننسى أن مستويات التقييم الواردة في الشكل رقم (١٩) تعتمد فقط على الوثائق التي يسترجعها النظام ، وليس على تلك التي لا يسترجعها ، أي أن الوثائق التي لم يتم التحقق منها هي :

١ - الوثائق التي تضاهي تسجيلاتها استراتيجية البحث ، ولكنها لم تسترجع (وكما سبق أن بينا فإننا لا يمكن أن نتوقع حدوث ذلك في النظم محكمة التصميم) .

٢ - الوثائق التي تضاهي الاستراتيجية المقصودة ، إلا أنها لم تسترجع .

٣ - الوثائق المتصلة بموضوع صيغة الطلب ، إلا أنها لم تسترجع .

٤ - الوثائق الصالحة لتلبية الحاجة إلى المعلومات ، إلا أنها لم تسترجع .

ويتطلب إصدار هذه الأحكام طريقة ما لتقدير عدد الوثائق التي تضاهي الطلب أو تتصل بموضوعه ، أو الصالحة لتلبية الحاجة إلى المعلومات ، وعجز النظام عن استرجاعها. وينبغي أن تكون هناك وسيلة ما للتحقق من بعضها على الأقل . ويتناول الفصل التاسع بعض الإجراءات المحتملة لتحقيق ذلك .

| العلاقة | السؤال الذي يمكن أن يوجه | المصطلحات | المؤهل للإجابة عن السؤال | أسباب الرغبة في الإجابة عن السؤال |
|---|--|---|---|--|
| ١ . بين استراتيجية البحث وبدائل الوثائق | هل تتفق المصطلحات الواردة في البديل ومنطق الاستراتيجية؟ | المضاهاة بين الاستراتيجية وبدائل الوثائق | أي اختصاصي معلومات متآلف مع النظام المستخدم | للتحقق مما إذا كانت برامج البحث أو إجراءاته تعمل بطريقة سليمة |
| ٢ . بين الوثيقة واستراتيجية البحث | هل هذه هي نوعية الوثائق، من الناحية الموضوعية ، التي أراد مسئول البحث استرجاعها؟ | المضاهاة بين الوثائق والاستراتيجية المقصودة | الشخص الذي يقوم بإعداد استراتيجية البحث، أي شخص آخر خبير في الموضوع والنظام | للتحقق من المشكلات الواردة في رقم ١ : التعرف على حالات الخطأ في الكشف، التحقق من المشكلات النظامية (العلاقات) في النظام والمشكلات الأخرى المتصلة باللغة ، والتحقق من الأخطاء التي وقعت في استراتيجية البحث |
| ٣ . بين الوثيقة وصيغة الطلب | هل هذه الوثيقة استجابة صحيحة للطلب المقدم للنظام؟ (هل موضوعها هو نفس الموضوع المطلوب؟) | اتصال الوثيقة بموضوع الطلب | أي شخص خبير في الموضوع أو مجموعة محكمين خبراء في الموضوع | للتحقق من المشكلات الواردة في رقم ٢ وكذلك التحقق من مشكلات الخطأ في تفسير الطلب من جانب مسئول البحث |
| ٤ . بين الوثيقة والحاجة إلى المعلومات | هل تسهم الوثيقة في تلبية حاجة المستفيد من المعلومات | صلاحية الوثيقة بالنسبة للحاجة إلى المعلومات | المستفيد دون سواء | للتحقق من المشكلات الواردة في رقم ٢ وكذلك التحقق من مشكلات تفاعل المستفيد مع النظام، والتي تؤدي إلى التناقض بين الحاجة إلى المعلومات التي أمكن التحقق منها والحاجة التي تم التعبير عنها (صيغة الطلب). |

الشكل رقم (١٩) : بعض العلاقات المهمة في استرجاع المعلومات .

الإنتاج الفكري حول الاتصال بالموضوع :

هناك بعض الأعمال الأخرى التي تتناول مشكلات الصلاحية والاتصال بالموضوع. ولا نقدم هنا مراجعة شاملة لهذا الإنتاج في هذا القسم ، وإنما نحيل القارئ إلى المصادر التي يمكن أن تساعد في توضيح بعض أوجه التمييز التي سبق أن أشرنا إليها ، والتي يمكن أن تقدم صورة أكثر اكتمالاً للعوامل التي تؤثر في قرارات الاتصال بالموضوع أو الصلاحية .

ويتبنى كمب (1974) Kemp وفوسكت (1972 ، 1970) Foskett التمييز بين الاتصال بالموضوع والصلاحية الذي التزمنا به في هذا الفصل . ويرى كمب أنه بالنسبة لبعض أغراض التقييم ، يمكن الاكتفاء بقرارات الاتصال بالموضوع ، إلا أنه بالنسبة لأغراض أخرى فإنه لا غنى عن قرارات الصلاحية . ويشير كمب إلى قرارات الاتصال بالموضوع باعتبارها عامة وموضوعية ، أما قرارات الصلاحية فخاصة وذاتية. ولا تتفق معه في ذلك تمام الاتفاق ؛ فالاتصال بالموضوع لا يتسم بالموضوعية ، وإذا كان يتسم بالموضوعية فعلاً لكان من الممكن أن يكون هناك اتفاق تام بين مجموعة المحكمين حول اتصال موضوع مختلف الوثائق بمختلف الطلبات . ومثل هذا الاتفاق التام غير محتمل إلى أبعد الحدود . كذلك يرى كمب أن هناك شيئاً من التناظر بين علاقة كل من الصلاحية والاتصال بالموضوع من جهة ، وتلك العلاقات التي تنطوي عليها ثنائيات المصطلحات التي تستعمل تبادلياً ، والمستمدة من مجالات أخرى ، مثل : **denotation** بمعنى الدلالة أو المعنى و **connotation** بمعنى الإيحاء أو المعنى ، و **semantics** بمعنى السيمية أو علم الدلالة و **pragmatics** بمعنى التداولية ، والاتصال الرسمي **formal** وغير الرسمي **informal** ، والمعرفة العامة والخاصة .

ويميز فوسكت (1972) Foskett بين الاتصال بالموضوع والصلاحية على نفس الأساس الذي التزمناه في هذا الفصل ، ويعرف الوثيقة المتصلة بالموضوع بالوثيقة التي «تنتمي إلى المجال / الموضوع / إطار الاتصال الذي تحدد معاملة مصطلحات الطلب ، كما استقرت بإجماع العاملين في هذا المجال » (P.77) أما الوثيقة الصالحة فهي التي «تضيف معلومات جديدة للرصيد المتوافر فعلاً في ذهن المستفيد ، والتي تفيده في العمل الذي دفعه لطلب المعلومات » (P.77) . كذلك يرى فوسكت أنه غالباً ، وليس دائماً ، ماتكون الوثائق الصالحة متصلة بالموضوع أيضاً ، والعكس بالعكس .

ويناقش كوبر (1971) Cooper قضية الاتصال بالموضوع بقدر كبير من الإسهاب. وهو يحدد في الأساس أوجه التمييز التي حددت هنا ، إلا أنه يستعمل مصطلح الاتصال المنطقي بالموضوع Logical relevance (أو الموضوعية topicality) بدلاً من الاتصال بالموضوع relevance ، والجدوى أو القابلية للاستثمار بدلاً من الصلاحية pertinence . ويرى كوبر في عمل آخر (1973) Cooper أنه ينبغي تقييم نظم استرجاع المعلومات بناء على جدوى نتائجها أو قابلية هذه النتائج للاستثمار : «إنها في الواقع الوثائق ذات القابلية المرتفعة للاستثمار ، وليس مجرد الوثائق المتصلة بالموضوع ، ما يريد المستفيد الاطلاع عليه » (P. 92) . ونتفق مع كوبر تمام الاتفاق في هذه النقطة .

ولم يفلح جوفمان (1964) Goffman في التمييز بين الاتصال بالموضوع والصلاحية ؛ فهو يعرف الاتصال بالموضوع بأنه « مقياس للمعلومات التي تحملها الوثيقة بالنسبة للاستفسار » (P. 201) ، وهو موقف مشابه لما ذهب إليه فوسكت وكيم ومؤلفا هذا الكتاب ، ويقابل ما يسميه كوبر « الاتصال المنطقي بالموضوع » . إلا أن جوفمان يستطرد قائلاً إن « أي مقياس للمعلومات ينبغي أن يعتمد على ما هو معروف فعلاً ، حقيقة لا ينبغي تجاهلها في أي تقدير لدى اتصال الوثيقة بموضوع الاستفسار » (P. 201) . وهذه العبارة قد اختلطت فيها الأمور إلى حد ما ، لأن التعبير « ما هو معروف فعلاً » يدل على علاقة الصلاحية لا علاقة الاتصال بالموضوع . وما يريد جوفمان تأكيده هو أنه لا يمكن الحكم على الاتصال بالموضوع بالنسبة لكل استفسار في مقابل وثيقة بمعزل عن بقية عناصر الموقف ، وإنما ينبغي الحكم على مدى اتصال الوثيقة بالموضوع ، بالنسبة للطلب في إطار القرارات الخاصة بالوثائق الأخرى بالنسبة لنفس الطلب . ونحن نقر الافتراض الذي يستند إليه هذا الرأي ، وهو أن الاتصال بالموضوع نسبي ويمكن الحكم عليه وفقاً لنوع ما من الموازين . وبعبارة أخرى ، فإنه ينبغي أن يكون من الممكن أن يطلب من محكمي الاتصال بالموضوع تقسيم مجموعة الوثائق إلى ثلاث مجموعات على الأقل :

- ١ - الوثائق التي تتصل ، بوضوح لا لبس فيه ، بموضوع صيغة طلب معين .
- ٢ - الوثائق المتصلة بموضوع صيغة الطلب ، ولكنها أقل اتصالاً من وثائق المجموعة الأولى .
- ٣ - وثائق لا تتصل بموضوع صيغة الطلب .

وبناء على ذلك ، فإن قرارات الاتصال بالموضوع ، تتسم بالنسبية ، لأنها تصنف الوثائق أو تقسمها إلى فئات تبعاً لمدي اتصالها بالموضوع . أما إذا كانت عبارة جوفمان تعني ضمناً أن الترتيب الذي تقدم به الوثائق للمحكمين يؤثر في قرارات الاتصال بالموضوع ، وأنه من الممكن للمحكم أن يقرر أن وثيقة بعينها لا تتصل بالموضوع لأنها تكرر وثيقة سبق أن رآها ، فإنها حينئذ لا تتفق والتعريفات الواردة في هذا الفصل . ولا يمكن لقرار من هذا النوع أن يكون قرار اتصال بالموضوع وإنما قرار صلاحية ، أي حكم قيمي صدر على الوثيقة بالنسبة لحاجة ما إلى المعلومات . وللأسف ، فإننا لا ندري بوضوح ما يقصده جوفمان فعلاً ، نظراً لأنه لم يفرق بين الاتصال بالموضوع من جهة والصلاحية من جهة أخرى ، كما أنه لم يفرق بين الاحتياجات الفعلية إلى المعلومات والاحتياجات التي أعرب عنها .

وقد قدم ولسون (1973) Wilson مصطلح الاتصال الظرفي بالموضوع Situational relevance ، والذي يعرفه « بالاتصال بالموضوع في موقف أو ظرف شخصي معين ، ولكن بالنسبة للموقف أو الظرف كما يراه هو ، لا كما يراه آخرون أو كما هو في «الواقع فعلاً» (P. 460) . كما يسجل أيضاً أن الاتصال الظرفي بالموضوع يتصل بالاهتمام الفعلي concern لا بمجرد الرغبة interest ؛ فمن الممكن للشخص أن يكون راغباً في شيء ما أو نشاط ما دون أن يكون مهتماً به ، أي دون أن يحفل أو يبالي بظروفه . ويرى ولسون أوعية المعلومات متصلة ظرفياً بالموضوع إذا كانت « تجيب أو تسهم في الإجابة عن الأسئلة التي تشغل البال فعلاً » (P. 463) . وتتوافق فكرة الاتصال الظرفي بالموضوع مع مجموعة العلاقات التي سبق التعرف عليها في هذا الفصل . وعلى الرغم من أن ولسون يضع المزيد من القيود على تعريف المصطلح ، فإن الاتصال الظرفي بالموضوع يدل على العلاقة بين المعلومات والحاجة المدركة إلى المعلومات ، لا على العلاقة بين المعلومات والحاجة الفعلية أو الحاجة المعرب عنها .

ويستعمل بلزر (1973) Belzer الاتصال بالموضوع حيثما استعملنا نحن الصلاحية : « فبعد أن يفرغ من قراءة الوثيقة ، يمكن للمستفيد أن يدرك على وجه التحديد ما إذا كانت صالحة بالنسبة له أم لا . ويحدث ذلك بصرف النظر عن السؤال الذي قدمه للنظام » (P. 301) . ويدرك بلزر بوضوح الفارق بين الاتصال بموضوع الطلب والصلاحية بالنسبة للحاجة إلى المعلومات . إلا أنه قد أثر تجاهل الأول .

وقد ناقش أوكونور (O'Connor (1967) قضية الاتصال بالموضوع في سلسلة من المقالات ، ويحلل في إحدى هذه المقالات تفصيلاً ، العلاقة بين الطلب والوثيقة ، وأثر وضوح الطلب على قرارات الاتصال بالموضوع . ونشر في مكان آخر تقريراً عن دراسة خبروية أو إمبريقية empirical (*) لمظاهر الاتفاق وأوجه الاختلاف في الحكم ما إذا كانت وثيقة معينة « تجيب عن » سؤال معين (O'Connor , 1969) . وعلى الرغم من أنه لم يفصح عن ذلك صراحة ، فإنه يعني ضمناً أن قرارات الاتصال بالموضوع تمثل أسساً لا غبار عليها لتقييم النظم ، أما قرارات الصلاحية فليست كذلك . وهو على الأقل ينتقد أولئك الذين يرون أنه ينبغي تقييم النظام بناء على احتياجات المستفيدين ، في الأساس ، على اعتبار أن « تلبية حاجة المستفيد » قلما تعرّف على وجه التحديد . وعلى الرغم من إمكان الاتفاق مع أوكونور حول النقطة الأخيرة ، فإن القول بأن تلبية حاجة المستفيد مبهمة أو غامضة إلى حد ما ، أو غير قابلة للتعريف ، لا يمكن قبوله . وكانت هذه النقاط موضوع مناظرة في الإنتاج الفكري ، من جانب كل من أوكونور (O'Connor (1967 a , b) ولانكستر (Lancaster (1968 b ، بما لم يرض أيًا منهما .

ويفرق سوانسون (Swanson (1977 بين الصلاحية بالنسبة لموضوع ما وصلاحية وعاء المعلومات بالنسبة لشخص ما لأنه يمدّه بمعرفة جديدة . ومن الصعب تبين كيف يختلف هذا عن التمييز بين الاتصال بالموضوع والصلاحية ، أو التمييز الذي يسجله في مقالة لاحقة ، بين الصلاحية « الموضوعية » والصلاحية الذاتية (Swanson , (1986) (**) .

هذا وقد حاول عدة باحثين استكشاف ما لمختلف أشكال بدائل الوثائق من أثر على قرارات الاتصال بالموضوع . وعادة ما تقارن هذه الدراسات بين قرارات الاتصال بالموضوع أو قرارات الصلاحية التي يتخذها المحكمون أنفسهم عندما تقدم لهم مستويات متفاوتة من المعلومات حول مجموعة معينة من الوثائق . فمن الممكن على سبيل المثال أن يطلب من المحكمين تسجيل توقعاتهم حول الاتصال بالموضوع بناء على العناوين ، ثم بناء على العناوين بالإضافة إلى المستخلصات ، أو العناوين بالإضافة إلى فقرات منتقاة من النص ، ثم يطلب منهم في النهاية الحكم على اتصال الوثائق نفسها بالموضوع .

(*) يستعمل هذا المصطلح عموماً للدلالة على تطبيق المنهج التجريبي في العلوم الاجتماعية . (المترجم) .

(**) استعمل سوانسون كلمة relevance دون pertinence . (المترجم) .

ثم تقارن النتائج للتعرف على مدى الاتفاق بين توقعات الاتصال بالموضوع التي تمت بناء على مختلف أشكال البدائل ، وقرارات الاتصال بالموضوع فعلاً بناء على الوثائق نفسها .

وقد نشر كل من راث ورفاقه (Rath, et al. (1961 ، ورزنك Resnick ، ودایم Dym (1967)، وكنت ورفاقه (Kent, et al (1967) ، وشايري وكورفيرست Shirey and (1967) Kurfeerst ، وساراسفك (Saracevic (1969 ، وماركوس وبننفلد وكوجل Mar- (1971) cus, Benenfeld and Kugel ، وبلزر (Belzer (1973 وغيرهم تقارير عن دراسات من هذا النوع . وكما يمكن توقعه ، فإن تنبؤات الاتصال بالموضوع تتحسن بوجه عام ، أي أن التنبؤات المبنية على البدائل تتفق مع القرارات المبنية على الوثائق ، تبعاً لتزايد المعلومات التي تتوافر للمحكمين . ويسمى ماركوس وبننفلد وكوجل (1971) نوعية البدائل ، بناء على ما لها من قيمة في تحقيق تنبؤات صحيحة بالاتصال بالموضوع ، بالقدرة الكشفية أو الدلالية indicativity . وقد تبين لهم أن القدرة الدلالية للتسجيلية تزداد بوجه عام، وبشكل مباشر تبعاً لطولها وفقاً لعدد الكلمات. وقد تبين لجينز (Janes (1991 ، في دراسة أحدث نسبياً ، أن المستخلصات كانت هي أفضل البدائل على الإطلاق في التنبؤ بالاتصال بالموضوع ، تليها العناوين ، ثم البيانات الوراقية ، والمصطلحات الكشفية . ويدعو الانخفاض الواضح في أداء المصطلحات الكشفية للاستغراب . ومما لا شك فيه أن مالقائمة المصطلحات الكشفية من قيمة كدليل على المحتوى ، تتوقف إلى حد بعيد على نوعية الكشف المطبق وعددا المصطلحات المستخدمة . ومن الممكن في بعض الأحيان للقوائم المطولة للمصطلحات الكشفية أن تقدم صورة للمحتوى الموضوعي للوثائق أكثر اكتمالاً من تلك التي يمكن أن تقدمها المستخلصات الكشفية الموجزة . وكما تبين لماركوس وبننفلد وكوجل (1971) فإنه إذا ما تساوت جميع الأمور الأخرى ، يصبح طول التسجيلية هو العامل الأساس الذي يحكم القدرة الدلالية (Lancaster, 1991) .

وقد تناولت سلسلتان رئيسيتان من الدراسات العوامل المؤثرة في قرارات الاتصال بالموضوع . وأولى هاتين السلسلتين أجرتها مؤسسة تطوير النظم System Development Corporation (SDC) ، ونشر كل من كوادرا وكاتر (Cuadra and Katter (1967 a, b) وكوادرا وكاتر وهولز وولاس (Cuadra, Katter , Holmes and Wallace (1967) تقارير

عنها . أما الثانية فقد أجرتها جامعة وسترن ريزيرف Western Reserve University ، ونشر كل من ريس (1966) Riss وريس وشولتز (1967) Rees and Schultz تقارير عنها.

وقد أمكن التحقق من عدد كبير من المتغيرات التي تؤثر في قرارات الاتصال بالموضوع ، ودراسة هذه المتغيرات في سلسلة دراسات مؤسسة تطوير النظم ، إلا أن الصلاحية بالنسبة للاحتياجات إلى المعلومات لم تحظ بالاهتمام . ويلخص كوادرا وكاتر (1967 b) Cuadra and Katter ما انتهت إليه هذه الدراسات على النحو التالي :

« تقدم الدراسات دليلاً لا يقبل الشك على أنه من الممكن لقرارات الاتصال بالموضوع أن تتأثر بكل من مهارات المحكمين واتجاهاتهم ، والوثائق ومجموعات الوثائق المستخدمة ، وصيغ التعبير عن الحاجة إلى المعلومات ، والتعليمات أو التوجيهات والظروف التي يتم فيها اتخاذ القرارات ، ومفاهيم وتعريفات الاتصال بالموضوع المتبعة في إصدار الأحكام ، ونوعية ميزان التقدير أو غيره من الوسائل المستخدمة للتعبير عن الأحكام . وتلقي هذه النتائج بظلال شك كثيفة على الأساس النظري لمعاملة مجاميع نقاط الاتصال بالموضوع ، كما يتم الحصول عليها عادة ، باعتبارها المعايير الملائمة تماماً لتقييم النظم أو النظم الفرعية » . (P. 98) .

وقد وضعت دراسات وسترن ريزيرف في اعتبارها أربعة متغيرات جوهرية تؤثر في قرارات الاتصال بالموضوع ، وهي :

(١) المرحلة التي يمر بها البحث (*) . (٢) الوثائق .

(٣) بدائل الوثائق . (٤) محكمو الاتصال بالموضوع .

وقد تمت هنا أيضاً دراسة الاتصال بالموضوع لا الصلاحية . وكما في دراسات مؤسسة تطوير النظم ، استطاع باحثو وسترن ريزيرف إثبات وجود عدد كبير من المتغيرات التي تؤثر في الحكم ما إذا كانت وثيقة معينة تتصل بموضوع طلب معين أم لا . فلدرجة دراية المحكم بالموضوع مثلاً أثرها في قرارات الاتصال بالموضوع ومدى الاطراد في اتخاذ مثل هذه القرارات من جانب مجموعة من المحكمين . كما أن من الممكن أيضاً إصدار أحكام اتصال بالموضوع بالنسبة لمجموعة معينة من الوثائق ومجموعة

(*) البحث هنا هو البحث العلمي أو المنهجي لا البحث عن الإنتاج الفكري في مرادف البيانات . (المترجم) .

معينة من صيغ الطلب ، تختلف إلى حد ما (الأحكام) تبعاً لاختلاف مراحل مشروع البحث، أي أنه من الممكن للوثائق التي يحكم أحد الباحثين بأنها أكثر ما تكون اتصالاً بالموضوع ، في بداية مشروع البحث ، ألا يصدر عليها الحكم نفسه على مشارف الانتهاء من المشروع ، عند تحليل نتائج البحث .

وقد وردت أكمل معالجة لقضية الاتصال بالموضوع في رسالة للدكتوراه قدمها ساراسفك (Saracevic (1970 a . كما أعد ساراسفك أيضاً ملخصاً موجزاً مفيداً جداً لأهم نتائج مختلف الدارسين المهتمين بإجراء التجارب في الاتصال بالموضوع ، خلال عشر سنوات (Saracevic , 1970 b) .

ونظراً للتحقق من عدد كبير جداً من العوامل المؤثرة في قرارات الاتصال بالموضوع، وخاصة في سلسلتي دراسات مؤسسة تطوير النظم ووسترن ريزيرف ، فإنه لا عجب على الإطلاق لإثارة شكوك خطيرة حول الحكمة في استناد محاولات تقييم النظم إلى قرارات الاتصال بالموضوع . إلا أنه ينبغي ألا ننسى أنه على الرغم من أن بعضاً من الجهود حول قرارات الاتصال بالموضوع تنطبق أيضاً وبالقدر نفسه على قرارات الصلاحية (كتأثير المرحلة التي يمر بها البحث والقدرة الدلالية للبدائل مثلاً) فإن كثيراً من هذه الجهود لا يتصل بالصلاحية اتصالاً مباشراً . فكثير من المتغيرات التي خضعت للدراسة بوجه خاص ، تتصل بمدى الاطراد في قرارات المحكمين ، ولا تؤثر بالضرورة في قرارات الصلاحية التي تتسم بالفردية إلى حد بعيد . وفي قرارات الصلاحية التي لا غنى عنها في تقييم مرافق المعلومات العاملة فعلاً ، ينصب الاهتمام على العوامل المؤثرة في اطراد قرارات المحكم نفسه ، لا بتلك المؤثرة بشكل مباشر في اطراد قرارات مجموعة المحكمين . وبعبارة أكثر تحديداً ، فإنه عند إجراء بحث في مرصد للبيانات لصالح مستفيد معين فإن بؤرة الاهتمام ينبغي أن تكون تقييم المستفيد للوثائق المسترجعة وفقاً لصلاحيتها بالنسبة لحاجته إلى المعلومات . كذلك ينبغي أن يكون مسئول البحث مهتماً بالعوامل المؤثرة في قرارات الصلاحية التي يتخذها المستفيد . إلا أن المرء لا يحتاج لأن يهتم ، وبشكل مباشر عندما يختلف مجموعة من الاختصاصيين الموضوعيين فيما بينهم حول أي الوثائق تتصل بموضوع الطلب ، عندما تعرض عليهم مجموعة الوثائق المسترجعة وصيغة طلب المستفيد . بل إن المرء قد لا يحفل بالعوامل المؤثرة في قراراتهم . وقد يكون لمثل هذه الدراسة أهميتها الأكاديمية ،

إلا أنها لا تتصل مباشرة بتقييم مرفق المعلومات هذا على وجه الخصوص . وكثير من الجهود الشيقة القيمة التي بذلت في دراسة مشكلة الاتصال بالموضوع لا تنطبق على مشكلة الصلاحية بشكل مباشر إلا في أضيق الحدود . والواقع أن معظم الدراسات لم تسهم مباشرة في الإحاطة بمشكلة الصلاحية ، لأنها أجريت في ظروف تجريبية منضبطة . أما دراسات الصلاحية فلا يمكن أن تتم إلا في سياق نظام معلومات معين يقدم خدماته لمستفيدين حقيقيين لديهم احتياجات حقيقية إلى المعلومات . ولا يمكن محاكاة هذا الموقف بنجاح في الظروف المختبرية .

ولا يعني ذلك أن مشكلة الاتصال بالموضوع ليست جديدة بالدراسة أو أن دراسات العوامل المؤثرة في قرارات الاتصال بالموضوع لا قيمة لها . والواقع أنه لبعض أغراض التقييم ، قد تكون قرارات الاتصال بالموضوع لا قرارات الصلاحية هي المطلوبة . إلا أن الخطر يكمن في التسليم بأن للدراسات الخاصة بالاتصال بالموضوع ارتباطها المباشر بالصلاحية ، والأهم من ذلك أن الأعداد الكبيرة من المتغيرات المؤثرة في الاطراد الذي يتم به إصدار أحكام الاتصال بالموضوع ، ترتبط مباشرة بالعوامل المؤثرة في قرارات الصلاحية ، وهي ليست كذلك فعلاً ، أو ليس من الضروري على الأقل أن تكون كذلك . ومما لا شك فيه أن كثيراً من العوامل تؤثر في قرار المستفيد ما إذا كانت وثيقة معينة صالحة بالنسبة لحاجته إلى المعلومات أم لا . وقد لا تكون هذه هي العوامل التي يمكن أن تؤثر في قرار المستفيد إذا ما طلب منه ، وبطريقة مصطنعة إلى حد ما ، أن يحكم ما إذا كانت الوثيقة نفسها متصلة بموضوع بيان أو صيغة طلبه . والحقيقة أنه على الرغم من أن كثيراً من الجهود قد بذلت حول الاتصال بالموضوع ، فقد كان التجاهل إلى حد بعيد من جانب الباحثين من نصيب المشكلة الأهم وهي الصلاحية ، وربما كان ذلك للسبب نفسه ، وهو عدم طواعيتها للدراسة التجريبية المنضبطة .

نقطة أخرى جديدة بالتركيز ، وهي أنه على الرغم من أننا قد لا نتوصل إلى اتفاق تام بين مجموعة من المحكمين حول أي الوثائق تتصل بموضوع بيان طلب معين ، فإن أوجه الاختلاف في حكم الاتصال بالموضوع ، ليس من الضروري أن تنال من قدرة أنواع معينة من التقييم . فقد بين كل من لسك وسالتون (1968) Lesk and Salton ، على سبيل المثال ، بوضوح لا لبس فيه أن عدم الاطراد في قرارات محكمي الاتصال بالموضوع قد لا يكون له من أثر على بعض الجوانب الداخلية لتقييم النظام . فإذا كان

هناك ، على سبيل المثال ، من يريد المقارنة بين ثلاثة سبل مختلفة لإجراء عمليات البحث في نظام معين ، فإنه من الممكن الخروج بالترتيب النسبي نفسه لكفاءة هذه البدائل (على أساس نسب الاستدعاء والتحقيق) ، أيأ كان المحكمون الخمسة مثلاً الذين يتخذون قرارات الاتصال بالموضوع ، أي أنه بعبارة أخرى ، يمكن أن يكون هناك عدم اطراد بين المحكمين ، إلا أن الترتيب النسبي لأساليب إجراء عمليات البحث قد لا يتغير . والواقع أن دراسة لسك وسالتون قد كشفت عن أنه حتى « الاختلافات الشاسعة في قرارات الاتصال بالموضوع » لم تسفر عن « مظاهر تباين جوهرية في متوسط الاستدعاء والتحقيق » بالنسبة لمختلف سبل إجراء عمليات البحث (P. 343) .

والولع بقضية الاتصال بالموضوع والصلاحية في تزايد مستمر نظراً لتزايد أعداد المكتبيين المهتمين بتقييم الفهارس المتاحة للجمهور على الخط المباشر . وقد تناول أوبراين (1990) O'Brien أهمية الاتصال بالموضوع في هذا السياق .

وقد نشر كل من شامبر وآيزنبرج ونيلان (Schamber, Eisenberg and Nilan 1990) مؤخراً مراجعة علمية ضافية إلى حد بعيد لقضية الاتصال بالموضوع ، وانتهوا إلى أن :

« ١ - الاتصال بالموضوع مفهوم معرفي متعدد الأبعاد ، يتوقف معناه إلى حد بعيد على إدراك المستفيدين للمعلومات والظروف التي يحتاجون فيها إلى المعلومات .

٢ - الاتصال بالموضوع مفهوم ديناميكي يتوقف على ما يصدره المستفيدون من أحكام على نوعية العلاقة بين المعلومات والحاجة إلى المعلومات في سياق زمني معين .

٣ - الاتصال بالموضوع مفهوم معقد ، إلا أنه نسقي أو منطقي وقابل للقياس إذا أمكن تناوله نظرياً وعملياً من وجهة نظر المستفيد » . (P. 774) .

وعلى الرغم من أن شامبر وآيزنبرج ونيلان قد قدموا إسهاماً مفيداً في الإنتاج الفكري في الموضوع (حتى على الرغم من أنهم يستعملون الاتصال بالموضوع حيثما يمكن لهذا الكتاب أن يستعمل الصلاحية) فإنهم يميلون للمبالغة في أهمية التوصل إلى تعريف يحظى بإجماع القبول للاتصال بالموضوع ، وقد بلغوا حد الادعاء بأن قالوا « إننا نعتبر تتبع أو ملاحقة تعريف الاتصال بالموضوع واحداً من أكثر تحديات علم المعلومات إثارة ومحرورية ، تحدٍ ينقلنا التغلب عليه إلى القرن الحادي والعشرين » . (P. 774) .

ونود أن نوّكد في الختام ، أنه بالنسبة لأغراض معينة للتقييم نحتاج إلى قرارات الاتصال بالموضوع ، وبالنسبة لأغراض أخرى فإن قرارات الصلاحية هي الأساس . وقد بذلت جهود كبيرة في الاهتمام بموضوع العوامل المؤثرة في الاتصال بالموضوع ، ولم تحظ العوامل المؤثرة في الصلاحية إلا بالنزر اليسير . والعوامل المؤثرة في الصلاحية ذاتية ووقتيّة ، إلا أن ذلك لا يقلل من أهميتها في تقييم أنشطة استرجاع المعلومات .

وتتناول الفصول من الرابع حتى الثامن مختلف العوامل المؤثرة في كفاءة نظم استرجاع المعلومات ، كما يصورها الشكل رقم (١٦) وهي على وجه التحديد التواصل بين المستفيد والوسيط ، والعوامل الخاصة بمرصد البيانات (التحليل الموضوعي والتعبير عن ناتج التحليل ومكونات اللغة) واختيار مرصد البيانات ، وعملية البحث .

الفصل الرابع

تواصل المستفيد والوسيط

لقد مهد الفصل الثالث لبعض المعايير التي يمكن بها تقييم خدمات استرجاع المعلومات ، كما أشار إلى أهم العوامل المؤثرة في كفاءة مثل هذه النظم . وتتناول الفصول من الرابع حتى الثامن عوامل الكفاءة هذه بشيء من التفصيل . ومن بين المشكلات التي واجهت المؤلفين تسلسل ترتيب هذه الفصول . وإذا ما اتبعنا التسلسل الزمني كما في الشكل رقم (١٦) فإننا يمكن أن نبدأ أولاً بالتفاعل بين المستفيد والوسيط، ثم ننتقل إلى اختيار مرصد البيانات ، ومناقشة استراتيجية البحث ، لننتهي بمعالجة العوامل الخاصة بمرصد البيانات (التكشيف والجوانب الأخرى للتحليل الموضوعي والأمور المتصلة باللغة) . وعلى الرغم من أن هذا تسلسل تتابعي منطقي ، إلا أنه ليس من الضروري أن يكون ترتيباً منطقياً للمعالجة ، لأن مناقشة اختيار مرصد البيانات واستراتيجية البحث ينبغي أن تستند إلى مناقشة مبكرة لعوامل مرصد البيانات المؤثرة في الكفاءة . فالتسلسل النهائي الذي وقع عليه الاختيار إذن وهو التفاعل بين المستفيد والوسيط ، والتحليل الموضوعي وإعداد بدائل الوثائق ، وقضايا اللغة ، واختيار مرصد البيانات وتقييمه ، واستراتيجية البحث ، بدأ بالنسبة لنا أكثر منطقية من التتابع الزمني.

الاحتياجات في مقابل الطلبات :

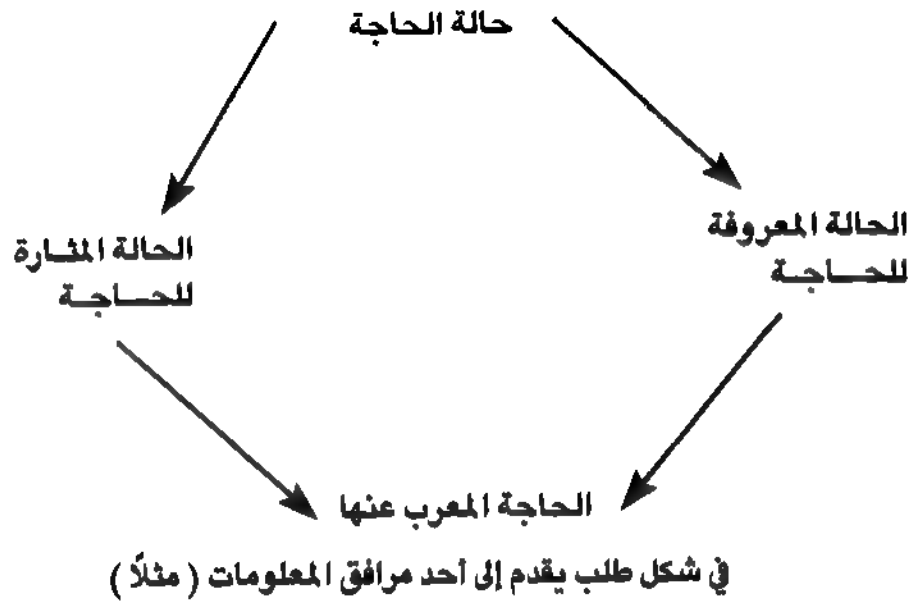
الخطوات الأولى في عملية استرجاع المعلومات (انظر الشكل رقم (١٦)) هي تلك التي يدرك فيها أحد أفراد مجتمع المستفيدين من خدمات أحد مرافق المعلومات ، الحاجة إلى

المعلومات، ويلجأ إلى المرفق في محاولة لتلبية هذه الحاجة . ويمكن للمستفيدين إجراء عمليات البحث الخاصة بهم أو تفويض المسؤولية إلى وسيط (اختصاصي معلومات) . ويتناول هذا الفصل البحث المفوض .

من المهم بمكان عند النظر في تقييم مرافق المعلومات ، التمييز بين احتياجات المجتمع المزمع خدمته والطلبات التي تقدم فعلاً للخدمة . ويمكن النظر إلى الاحتياجات باعتبارها أكثر تعدداً من الطلبات (الاحتياجات المعرب عنها) نظراً لأن الاحتياجات إلى المعلومات لا تتحول جميعها إلى طلبات . ويتعين على مديري مرافق المعلومات الاهتمام بالتعرف على احتياجات المجتمع المزمع خدمته إلى المعلومات ، فضلاً عن التحقق من مظاهر التناقض بين الاحتياجات والطلبات . ولا شك أن من بين الجوانب المهمة للتقييم التحقق مما بين الاحتياجات والطلبات من اختلافات وفقاً لاعتبارات كمية (كم عدد الاحتياجات التي لم تتحول إلى طلبات) واعتبارات نوعية (ما هي أنواع الاحتياجات التي لم تتحول إلى طلبات ، وما هي العوامل التي تقرر ما إذا كانت الحاجة تتحول إلى طلب أم لا ، وإلى أي مدى تعبر طلبات المستفيدين بدقة عن احتياجاتهم الفعلية إلى المعلومات) . وتكاد معظم عمليات تقييم مرافق المعلومات تقريباً تركز ، للأسف على مجرد قياس مدى قدرة مرفق المعلومات على تلبية طلبات المستفيدين . وهذا مدخل سطحي إلى حد ما للتقييم ، حيث يتجاهل الاحتياجات التي لم يعرب عنها المستفيدون ، كما يفترض أن ما يقدمه المستفيدون من طلبات يتطابق مع احتياجاتهم ، وهذا افتراض محفوف بالمخاطر .

والتركيز على الطلبات التي تقدم فعلاً ، دون غيرها ، وقبولها بقيمتها الظاهرية أشبه ما يكون بالتركيز على النتوء الظاهر لجبل جليدي ضخم ، وتصور تمثيل هذا النتوء للكتلة الضخمة القابعة تحته تمثيلاً كاملاً . وقد بين لاين (1973) Line مخاطر هذا المدخل . ومما لا شك فيه أنه من الممكن التعرف على الطلبات بشكل أيسر بكثير من التعرف على الاحتياجات . إلا أنه لا يمكن تجاهل جانب مهم من جوانب التقييم ، لا شيء وإنما لأنه من الصعب إدراكه .

وقد حاول كوشان (1979) Kochen التمييز بين الاحتياجات والمشكلات والتعبير عن الاحتياجات والمشكلات . ومن الممكن وبشكل غاية في البساطة تصوير الموقف على النحو التالي :



ويميز كوشان بين حالة الحاجة والتعرف على الحاجة أو إثارتها ، ومن الممكن لبعض الاحتياجات أن تثار دون التعرف عليها ، كما يحدث في بعض الأحيان أن يكون من الممكن التعرف على الحاجة دون إثارتها . ومن المهم بمكان في سياق مرافق المعلومات التمييز بين الاحتياجات إلى المعلومات ، والتعرف على الاحتياجات والتعبير عنها . ولا يمكن لنظم المعلومات الاستجابة لاحتياجات الأفراد إلى المعلومات كاحتياجات ، وإنما تستجيب فقط لإعرابهم عن احتياجاتهم ، ويعني ذلك أنه يتعين على من يحتاج إلى المعلومات أن يتعرف على تلك الحاجة ، كما ينبغي أن يكون لديه الدافع الكافي لاتخاذ خطوات تلبية لها . ولا يتم الإعراب عن الحاجة في شكل طلب يقدم إلى أحد مراكز المعلومات ، إلا بعد التحقق منها ويتوافر لدى المستفيد الدافع الكافي . وعلى مدى قدرة المستفيد على التحقق من الطبيعة الحقيقية للحاجة إلى المعلومات ، وعلى درجة الدقة في تصوير هذه الحاجة في التعبير عنها ، في بيان الطلب ، يتوقف وإلى حد بعيد نجاح مرافق المعلومات في تلبية تلك الحاجة ، حيث لا يمكن لمرافق المعلومات أن يعمل إلا بناء على الطلب المقدم (الحاجة المعرب عنها) ، كما أنه لا يستطيع الاستجابة لاحتياجات لم يتم التحقق منها ، أو احتياجات تم التحقق منها ولم يتم الإعراب عنها . ومن أصعب التحديات التي تواجه أي مرافق للمعلومات يعمل بطريقة البحث المفوض ، التأكد من دقة مطابقة الاحتياجات المعرب عنها للاحتياجات التي تم التحقق منها . فليس من السهل دائماً على من يحتاج إلى المعلومات أن يعرب عن تلك الحاجة بوضوح لا لبس فيه ، لمن يتولى مهمة البحث عن المعلومات .

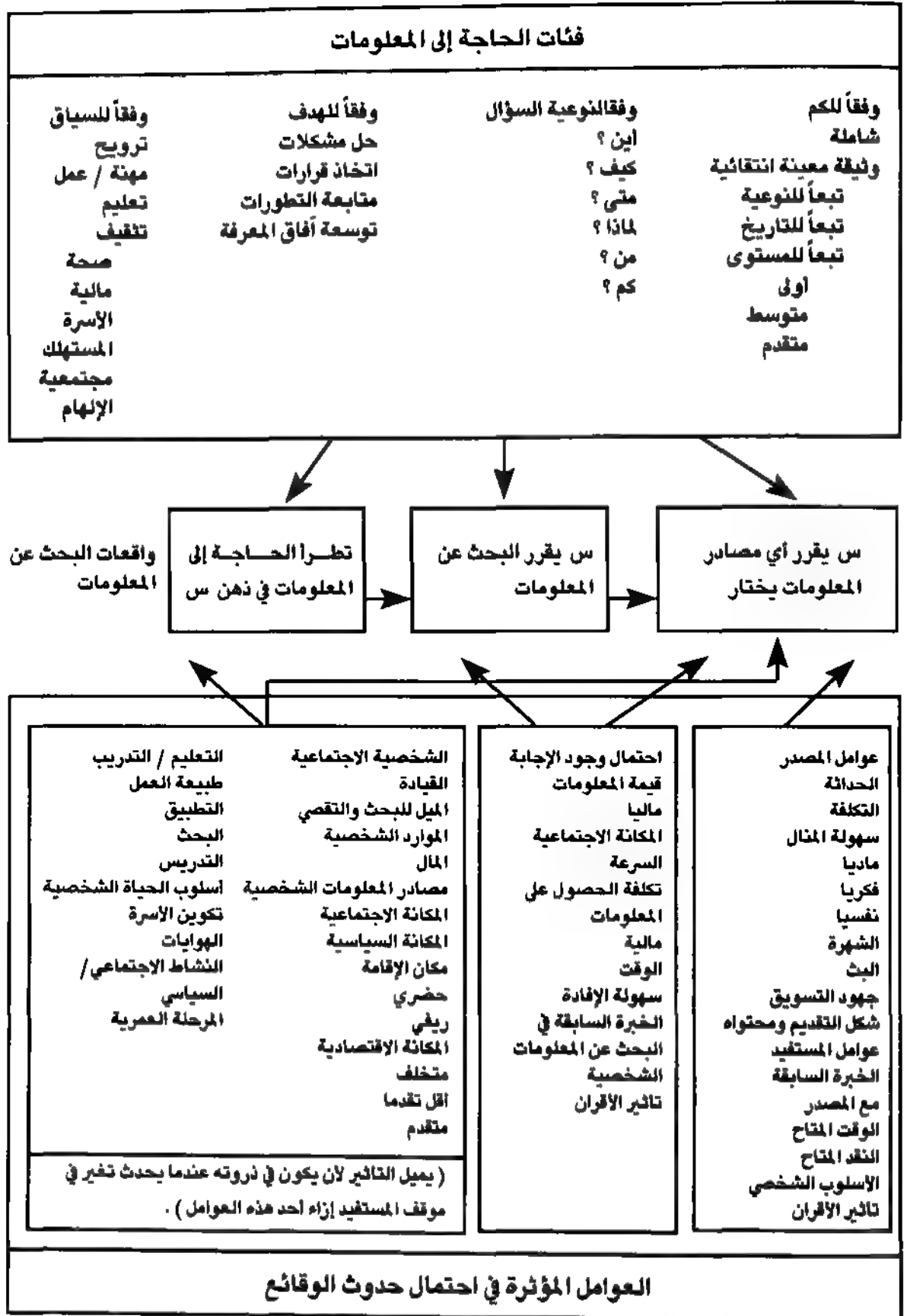
والتمييز بين الحاجة إلى خدمات المعلومات والطلب على هذه الخدمات أمر مهم ، وعلى مديري مرافق المعلومات أن يهتموا بتقييم ما يقدمون من خدمات بناء على مدى مطابقة هذه الخدمات لاحتياجات المستفيدين المحتملين قدر اهتمامهم بتقييمها بالنسبة لما يقدمه المستفيدون الفعليون من طلبات . وينطوي قصر اعتبارات التقييم على الطلبات التي تقدم في الظروف الآنية لمرافق المعلومات ، على تجاهل لاحتياجات المستفيدين الحاليين التي لا تتحول إلى طلبات ، واحتياجات من لا يطلبون الخدمات حالياً ، ممن ينتمون إلى المجتمع المزمع خدمته .

ويحدث في كثير من الأحيان أن يفوق عدد غير المستفيدين المحتملين من الخدمة أعداد المستفيدين الفعليين بمراحل . هذا بالإضافة إلى أن احتياجات المستفيدين الفعليين إلى المعلومات لا تتحول جميعها إلى طلبات . وهناك مشكلة أخرى ، وهي أن ما يقدم إلى المرفق من طلبات ليست جميعها تعبيراً كاملاً عن الحاجة إلى المعلومات الكامنة وراء الطلب . وأحياناً ما يطلب المستفيدون أقل مما يحتاجون ، وهناك ميل إلى حد ما من جانب المستفيدين من مرافق المعلومات في الواقع ، لطلب ما يظنون أن بإمكان النظام تقديمه لا ما يريدونه فعلاً (Lancaster, 1968 a) . وغالباً ما يعني ذلك أن الطلب أعم بكثير من الحاجة إلى المعلومات التي دفعت إليه .

هذا ، ويصور الشكل رقم (٢٠) الفئات الأساسية للحاجة إلى المعلومات ، وفقاً للسياق ، والغرض ، ونوع السؤال ، ومقدار المعلومات التي تدعو الحاجة إليها ، وأهم العوامل المؤثرة في احتمالات أن :

- (١) تنشأ الحاجة إلى المعلومات في ذهن شخص ما .
- (٢) يقرر الشخص التماس المعلومات التي يحتاج إليها .
- (٣) اختيار مصدر بعينه للمعلومات .

ويتسم الشكل بالعمومية ، حيث يتناول جميع أنماط الحاجة إلى المعلومات ، بما في ذلك احتياجات الجمهور العام ، وربما احتاج إلى تعديل لكي ينطبق بمزيد من الوضوح على بيئة بعينها أو شخص معين (كالحاجة إلى المعلومات مثلاً في إحدى الشركات الصيدلية ، أو احتياجات الباحث الكيميائي إلى المعلومات) وإن كان من الممكن لكثير من التأثيرات أن تظل محتفظة بفعاليتها .



الشكل رقم (٢٠) : تعقد موقف « الحاجة إلى المعلومات » .

ومعظم التأثيرات الواردة في الشكل تفسر نفسها بنفسها ، وإن كان من الممكن لعدد قليل منها أن يتطلب توضيحاً . ويتوقف تحول حاجة معينة إلى المعلومات إلى طلب يقدم لمرفق المعلومات ، إلى حد بعيد على قيمة حل مشكلة المعلومات ؛ فالحل الذي تبلغ قيمته ١٠٠,٠٠٠ دولار بالنسبة لشركة معينة ، تزيد احتمالات تحوله إلى طلب للمعلومات عن احتمالات حل تبلغ قيمته ٥٠٠ دولار أو حل لا تبدو له قيمة مالية على الإطلاق ، كما يتوقف تحول الحاجة إلى طلب أيضاً على مدى احتمال وجود الحل في الإنتاج الفكري . ولا ينبغي النظر إلى قيمة المعلومات على أسس مالية فقط ؛ فيحدث في بعض الأحيان أن تدعو الحاجة إلى المعلومات مثلاً لإثبات حقيقة أو اكتساب مكانة . ومما لا شك فيه أن كثيراً من التأثيرات الواردة في الشكل تتصل اتصالاً وثيقاً ببعضها البعض ، كما تتبادل المواقع فيما بينها . ويمكن لتكلفة خدمة المعلومات أن تكون اعتباراً تافهاً أو جديراً بالإهمال إذا كانت القيمة التي أمكن إدراكها لحل المشكلة عظيمة .

ولا تدخل العوامل المتصلة باحتمال نشوء الحاجة إلى المعلومات ، أو احتمال البحث عن حل ، بالطبع ، تحت السيطرة المباشرة لمديري مرافق المعلومات . أضف إلى ذلك أنه من المهم التحقق من وجود هذه العوامل وأن لها أثراً لا يمكن إنكاره على الحاجة إلى المعلومات أو اللجوء إلى مرافق المعلومات . وتقع عوامل النظام (المصدر) أكثر من غيرها مباشرة تحت سيطرة مديري المرافق ، إلا أن هذه العوامل تتأثر بقوة أيضاً بالمؤثرات الخارجية ، كالتطورات التقنية والتكاليف الخارجية ، والتي لا تخضع مباشرة لسيطرة مرافق المعلومات .

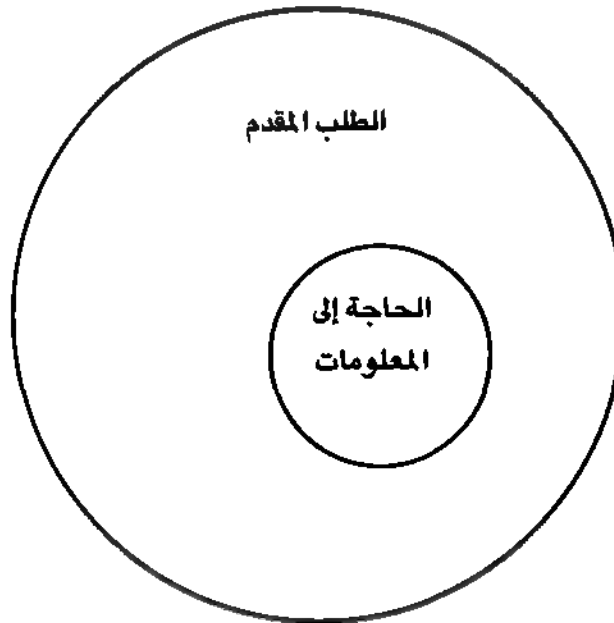
ولا يمكن لمركز المعلومات أن يؤثر مباشرة في الحاجة إلى المعلومات في نطاق المجتمع الذي يعمل على خدمته ، إلا أنه من الممكن ولا شك أن يؤثر في الطلب على خدمة المعلومات ، من خلال التعرف على عوامل النظام التي تؤثر في الطلب ، ومن خلال التقييم المستمر لمدى مطابقة الخدمات المقدمة للاحتياجات القائمة في المجتمع ، ومن خلال تقييم مدى السرعة والدقة والاكتمال في تلبية الطلبات .

ومن الممكن للطلبات التي تقدم للمرفق الرسمي للمعلومات أن تكون متصلة باحتياجات المعلومات التي لم يستطع المستفيد تلبيتها من مصدر آخر . فمن الأرجح أن يكون المستفيد قد حاول الحصول على المعلومات من مصدر أقرب منالاً ، كالمجموعة

الشخصية من الإنتاج الفكري أو شخص آخر . وعلى ذلك ، فإن الطلبات التي ترد إلى مرفق المعلومات يمكن أن يكون من الصعب لا من اليسير تلبيتها .

طلبات المعلومات :

تناولنا حتى الآن في هذا الفصل بعض العوامل التي تقرر ما إذا كان من الممكن للفرد أن يحتاج إلى المعلومات ، وما إذا كان من الممكن أن يعمل على تلبيتها . وهناك أمر آخر مهم ينبغي أخذه في الحسبان ، وهو العوامل التي تؤثر فيما إذا كان الطلب المقدم إلى مرفق المعلومات (الحاجة المعرب عنها) يعبر بدقة عن حاجة المستفيد الحقيقية إلى المعلومات .



الشكل رقم (٢١) : الطلب أعم من الحاجة إلى المعلومات . معظم الوثائق المسترجعة لا تتصل بالموضوع .

نفترض أنه يتعين على الشخص الذي يحتاج إلى المعلومات أن ينقل رغبته هذه إلى أحد العاملين بمركز المعلومات بواسطة الهاتف أو البريد أو الزيارة الشخصية ، فمما لا شك فيه أن لهذا التفاعل بين المستفيد والوسيط أهميته البالغة بالنسبة لعملية استرجاع المعلومات برمتها . ويمكن للغة النظام أن تكون ملائمة للتعبير عن المفاهيم الواردة في الطلب ، كما يمكن لاستراتيجية البحث أن تكون تعبيراً مكتملاً دقيقاً عن الطلب ، كما يمكن لتكشيف مرصد البيانات أن يكون كاملاً ودقيقاً ومطرداً . إلا أن كل

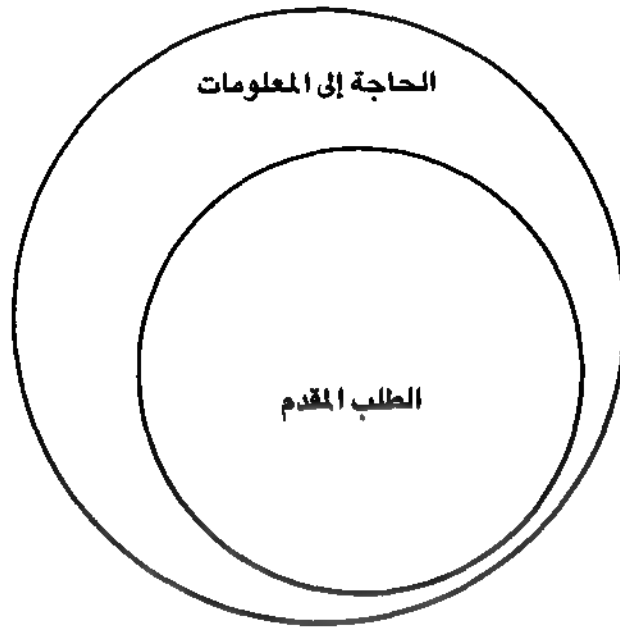
هذه الأمور لا قيمة لها بالنسبة لمستفيد معين إذا كان الطلب (الحاجة المعرب عنها) تعبيراً غير ملائم عن حاجة المستفيد الحقيقية .

ولكي يكون البحث في نظام الاسترجاع ناجحاً ، فإن الطلب المقدم ينبغي أن يكون تقديراً تقريبياً معقولاً للحاجة إلى المعلومات . وكلما زاد التباين بين الطلب المقدم والحاجة إلى المعلومات انخفضت احتمالات نجاح البحث . وليس من السهل دائماً على المستفيد ، للأسف أن يصف حاجته إلى المعلومات ، بشكل مكتمل دقيق ، لشخص آخر . وفي تقييم لانكستر (1968 a) Lancaster لمرصّد بيانات المدلرز MEDLARS ، كانت نسبة مئوية كبيرة من الأخطاء في ٣٠٠ عملية بحث ترجع إلى التفاعل غير الملائم بين المستفيد والوسيط ، مما أسفر عن طلبات شفوية ، لم تكن تعبر عن احتياجات المستفيدين الحقيقية إلى المعلومات بشكل ملائم .

وعادة ما يميل المستفيدون إلى تقديم طلبات أعم بكثير من الحاجة الفعلية إلى المعلومات ، وربما كان ذلك لأن المستفيد يظن أن النظام يمكن أن يعمل على مستوى أعرض لا على مستوى أكثر تخصيصاً (الشكل رقم (٢١)) . وبعبارة أخرى ، فإن المستفيد يسلك هنا بناء على توقعاته من النظام وما يمكن أن يقدمه . ومن المؤكد أن هذا هو التفسير الوحيد لطلب الباحث العلمي المعروف على المستوى الوطني في مجال السرطان ، إجراء بحث عن السرطان في الجنين أو الطفل حديث الولادة ، وهو موضوع يحظى بوفرة في الإنتاج الفكري ، الذي استرجعه النظام ، وكان كله تقريباً لا يتصل بحاجة العالم الدقيقة إلى المعلومات (1968 a) Lancaster . فما كان يبحث عنه فعلاً هو المعلومات عن موضوع دقيق جداً ، يتوافر فيه قدر ضئيل جداً من الإنتاج الفكري ، وهو على وجه التحديد العلاقة بين العيوب الخلقية والسرطان على المستوى الخلوي . وكانت الوثائق المسترجعة تضاهي الطلب الذي تقدم به تماماً ، ولكنها من المؤكد لم تكن صالحة بالنسبة لحاجته إلى المعلومات .

ونادراً ما يكون الطلب أكثر تخصيصاً من الحاجة إلى المعلومات (الشكل رقم (٢٢)) . وكمثال أمامنا المستفيد الذي طلب المعلومات عن مرور الأحماض الدهنية عبر الحاجز المشيمي وعن المستويات الطبيعية للأحماض الدهنية في المشيمة أو الجنين . وبالمراجعة تبين أن هذا الباحث كان يهتم فعلاً بموضوع أعرض ، وهو على وجه التحديد

مرور الدهون (أي المواد الدهنية بوجه عام) عبر الحاجز المشيمي والمستويات الطبيعية للدهون في المشيمة أو الجنين أو الطفل حديث الولادة . وعادة ما يؤدي الطلب الأكثر تخصيصاً من الحاجة الفعلية إلى المعلومات ، إلى العجز عن استرجاع بعض الوثائق التي كان من الممكن أن تكون لها قيمتها بالنسبة للمستفيد . وهذا الموقف أكثر تعقيداً من الموقف المقابل ، نظراً لأن التوسع في الطلب لا يمكن أن يتحقق إلا نتيجة للتصفح والبحث الإيعازي ، ويمكن ألا يحدث على الإطلاق عندما يعهد الشخص الذي يحتاج إلى المعلومات بمهمة إجراء عملية البحث لشخص آخر . وفي المثال السابق كان الطلب المقدم يعبر عما كان العالم يظن أنه يريده عندما لجأ إلى النظام في البداية . ولم يدرك أن طلبه كان محدوداً جداً ، وأن هذه الوثائق التي لا تدخل في مجال طلبه المقدم كانت مفيدة أيضاً في الحقيقة ، إلا بعد أن رأى بعض الوثائق التي لا تدخل في صميم طلبه المقدم ، حيث تتناول على وجه التحديد تبادل الدهون بين الأم والجنين ، ومستويات الدهون في حديثي الولادة .



الشكل رقم (٢٢) : الطلب أكثر تخصيصاً من الحاجة إلى المعلومات : لا يتم استرجاع بعض الوثائق الصالحة .

نوعية طلبات المستفيدين :

يكن أحد المصادر الأساسية للخطأ في أي نظام للمعلومات يعمل وفقاً للبحث المفوض، في التواصل بين المستفيد والنظام ، حيث يتعين على المستفيد أن يعبر عن

احتياجاته إلى المعلومات ، بشكل مكتمل دقيق لشخص آخر . وقد حاول عدد من الباحثين التحقق من العمليات التي ينطوي عليها تحويل الحاجة الكامنة أو المستترة إلى المعلومات إلى طلب يقدم إلى أحد مرافق المعلومات . وقد تعرف تيلور (Taylor (1967 على أربعة مستويات للحاجة :

١- الحاجة الكامنة . ٢- الحاجة الواعية .

٣- الحاجة التي تحدد شكلها . ٤- الحاجة القابلة للمداولة أو الحل الوسط .

فعل المستوى الكامن عادة ما تكون الحاجة مبهمة أو ربما تكون كامنة في اللا شعور؛ فهناك مشكلة ينبغي حلها أو قرار ينبغي اتخاذه ، إلا أن الشخص لم يصبح قادراً بعد على تحديد المعلومات التي يمكن أن تكون مفيدة . أما في المستوى المدرك أو الواعي فإن وصف الحاجة يكون في ذهن الفرد إلا أنه قد لا يكون قد تحددت معالمه بعد . أما على مستوى الشكل المحدد فإن الحاجة تكون قد تم التعرف عليها بقدر كافٍ من الوضوح إلا أنه لا يمكن الإفصاح عنها . أما الحاجة القابلة للتسوية أو الحل الوسط فهي الطلب الذي يقدم فعلاً إلى مرفق المعلومات ، وهي قابلة للتسوية لأنه من الممكن التعبير عنها بشكل يعتبره المستفيد قابلاً للفهم من جانب مرفق المعلومات إلا أنه لا يعبر بدقة عن الحاجة الحقيقية .

وينظر بلكن (Belkin (1980 وبلكن وآخرون (Belkin et al . (1982 a, b إلى مشكلة استرجاع المعلومات باعتبارها مشكلة مضاهاة حالة معرفية غير سوية (ما لا يعرفه السائل أو المستفيد) مقابل حالة معرفية مترابطة منطقياً (ما يعرفه المؤلفون كما تتضمنه أعمالهم المنشورة) . وهم يرون أن الاحتياجات إلى المعلومات غير قابلة للتوصيف الدقيق ، وإنما يمكن أن نكون قادرين على أن نحصل من المستفيدين على «بيان بالمشكلة» يمكن منه الخروج بتصوير للحالة المعرفية غير السوية (التي يرمز إليها بالأحرف الثلاثة ASK) anomalous state of knowledge الكامنة وراء الحاجة .

هذا ، ويرى آلن (Allen (1988 أن الطريقة التي يدرك بها المستفيدون احتياجاتهم ويعبرون عنها تتأثر بالبنى المعرفية (الخطط Schemata) التي ينظمون بها معرفتهم بالموضوع ، وبالخطط التي ترد فيما يوجهه الوسطاء من أسئلة . ويمكن للاختلافات في التصورات الشخصية أو الأطر المرجعية أن تعني أنه من الممكن للمشكلة نفسها أن تؤدي

إلى إثارة أسئلة مختلفة تمام الاختلاف في أذهان مختلف الأفراد . ومن المشكلات الجوهرية فعلاً في التفاعل بين المستفيد والوسيط احتمال التباين بين التصورات الشخصية (الإطار المرجعي) للمستفيد وتلك الخاصة باختصاصي المعلومات . أضف إلى ذلك أنه من الممكن أيضاً لمختلف « التصورات أو الخطط الخارجية » (كاللغة المستخدمة في كشف مرصد البيانات مثلاً) أن يكون لها أثرها القوي على التفاعل بين المستفيد والوسيط .

وينظر تيلور (1967) Taylor إلى التفاعل بين المستفيد والوسيط باعتباره ينطوي على تجهيز السؤال عن طريق سلسلة من المصافي أو المرشحات التي يعمل كل واحد منها على تنقية إدراك الوسيط لما يريده المستفيد فعلاً . ومن الممكن تحديد مرشحات تيلور على النحو التالي :

- (١) تعبير المستفيد عن الحاجة .
- (٢) إدراك الوسيط لأهداف المستفيد ودوافعه .
- (٣) السمات الشخصية للمستفيد .
- (٤) معرفة الوسيط بمحتوى وبنية مرصد البيانات الذي يمكن البحث فيه .
- (٥) إدراك الوسيط لنوعية الإجابة المطلوبة (حقيقة محددة بعينها ، أو عدد محدود من الإشارات الوراقية الصالحة أو بحث شامل مثلاً) .

- ١ - قدرة المستفيد على تحديد الحاجة إلى المعلومات في ذهنه .
- ٢ - قدرة المستفيد على التعبير عن نفسه .
- ٣ - مهارة الوسيط كهمزة وصل .
- ٤ - درجة توافق البنية المعرفية الشخصية (الإطار المرجعي) للوسيط مع تلك الخاصة بالمستفيد .
- ٥ - توقعات المستفيد حول قدرات النظام .
- ٦ - مدى تعقد الموضوع المتداول (المجرد في مقابل المحسوس مثلاً) ومدى إمكان التعبير عنه بدقة .
- ٧ - الأدوات المساعدة للتفاعل التي يقدمها النظام (نماذج طلب البحث مثلاً) .

الشكل رقم (٢٣) : العوامل المؤثرة في نوعية التفاعل بين المستفيد والوسيط .

وهناك الآن قطاع كبير من الإنتاج الفكري حول المقابلة التي تتم بين المستفيد واختصاصي المعلومات (مثل كروم Crum, 1969 وإنجفرسن Ingwersen, 1982 وكاتس Katz, 1987 وكنج King, 1972 وناب Knapp, 1976 ولينش Lynch, 1978 وماركي Markey, 1981 ورولوف Roloff, 1979 وسومرفيل Somerville, 1977 وهوايت Somerville, 1977 وWhite, 1981, 1983, 1985) ، وغالباً ما يتم التمييز بين خصائص ونوعية الأسئلة «المفتوحة» في مقابل الأسئلة «المغلقة» في المقابلة . والأسئلة المغلقة هي تلك التي يمكن الإجابة عنها بنعم أو لا ، بينما تتطلب الأسئلة المفتوحة إجابات أطول من ذلك . وتميل الأسئلة المفتوحة للكشف عن المزيد حول سياق المستفيد ، بينما تميل الأسئلة المغلقة للتقيد بالطلب الأصلي .

وينبغي أن يكون واضحاً أن لكثير من العوامل أهميتها في تحديد نوعية التفاعل بين المستفيد والوسيط (الشكل رقم (٢٣)) . وقد ألحنا إلى عدد من هذه العوامل فعلاً ، أما العوامل الأخرى فواضحة في حد ذاتها .

طرق تفاعل المستفيد والوسيط :

يمكن لبعض طرق التفاعل بين المستفيد والوسيط أن تكون أكثر فعالية من غيرها . ويمكن أن نفترض أنه عندما يزور المستفيد مركز المعلومات بشخصه ويناقش حاجته إلى المعلومات مع أحد اختصاصيي المعلومات ، فإنه من المحتمل أن يكون بيان الطلب الناتج عن ذلك تصويراً للحاجة إلى المعلومات أفضل من الطلب الذي يرد إلى مركز المعلومات بالبريد ، دون الإفادة من التفاعل المباشر بين المستفيد واختصاصي المعلومات .

ومن العجيب أن الدليل المستخلص من دراسة مرصد بيانات المدلرز MEDLARS (Lancaster, 1968 a) لا يدعم هذا الفرض ، وإنما العكس تماماً في الحقيقة . فإذا أخذنا عمليات البحث الثلاثمئة التي استند إليها تقييم المدلرز ، وقسمناها إلى مجموعتين ، الأولى تتكون من عمليات البحث المعتمدة على طلبات مقدمة في الزيارة الشخصية لمركز المعلومات ، والثانية من عمليات البحث التي تمت بناء على طلبات مقدمة بالبريد ، فإن كفاءة المجموعة الثانية تبدو أفضل بوضوح؛ أي أن عمليات البحث التي أجريت بناء على الطلبات الواردة بالبريد كانت قادرة في المتوسط على استرجاع عدد كبير من الوثائق التي أقر المستفيدون صلاحيتها وعدد قليل من الوثائق التي أقر عدم صلاحيتها .

وكان هذا الاكتشاف غير متوقع من جانب الغالبية العظمى من المهتمين بالقضية، وهو يدل على أن المستفيدين كانوا قادرين على إيصال احتياجاتهم من المعلومات تحريراً بشكل أكثر فعالية من التعبير الشفوي. ثم بدأت التحليلات التي أجريت لهذه النتائج بعد ذلك في الكشف عن بعض أسباب هذا الموقف. فيبدو أن المستفيدين الذين يكتبون طلباتهم على نموذج الطلب أو في خطاب، في خلوة بمنازلهم أو بمكاتبهم، يتمتعون عموماً بميزتين لا تتوافران للمستفيدين الذين يزورون مركز المعلومات بأشخاصهم. الميزة الأولى أن الحاجة إلى كتابة طلب تفرض الانضباط على المستفيدين؛ فهم مضطرون للتفكير فيما يبحثون عنه فعلاً، ويتعين عليهم أن يحاولوا التعبير عنه كتابة. والميزة الثانية أنهم لا يتأثرون بقيود النظام؛ فنظراً لبعدهم المكاني عن النظام فإنهم لا يميلون للنظر إليه باعتباره يمكن أن يكون به بعض مظاهر القصور في لغته أو في قدرته على البحث. ويميل المستفيدون في ظل هذه الظروف لوصف حاجتهم إلى المعلومات بلغتهم، ويطلبون ما يريدونه فعلاً، لا ما يظنون أن بإمكان النظام تقديمه. وتميل الطلبات التي تقدم إلى مراكز المعلومات كتابةً لأن تكون وصفاً دقيقاً بشكل معقول للمعلومات التي يتم البحث عنها فعلاً.

وفي مقابل ذلك انظر إلى أولئك المستفيدين الذين يزورون مراكز المعلومات بأشخاصهم؛ فلا يمكن أن يكونوا قد مروا بحالة الانضباط العقلي المصاحب لكتابة بيان الطلب، ومن ثم فإنهم قد لا يكونون قد كونوا فكرة كاملة عما يريدونه فعلاً. وعندما يأتون إلى مركز المعلومات ويحاولون وصف احتياجاتهم فإن هناك احتمالاً قوياً جداً لأن يكونوا متأثرين، وربما بلا وعي منهم، بقيود النظام. وفي ظل هذه الظروف فإنه من المحتمل جداً أن يسألوا عما يظنون أن بإمكان النظام أن يعطيهم، لا عن المعلومات المحددة التي يبحثون عنها فعلاً. وهذه الظاهرة معروفة جداً للمكتبيين. ومن أمثلتها المستفيد الذي يأتي إلى المكتبة بحثاً عن عنوان فندق معين في كوبنهاجن، إلا أنه بدلاً من السؤال عن هذا العنوان يطلب الاطلاع على الكتب حول الارتحال في الدول الاسكندنافية.

ورغم نواياه الطيبة، يميل اختصاصي المعلومات للتأثير على المستفيد في الطريقة التي يصف بها هذا الأخير حاجته. ويمكن لهذا التأثير أن يكون ضاراً في حالة ما إذا كان المستفيد لم يتدبر حاجته إلى المعلومات ولم يفصح عنها كتابة. وتبلغ احتمالات

التحريف ذروتها عندما يناقش المستفيد اختصاصي المعلومات الطلب بناء على اللغة المقيدة للنظام، حيث يضع ذلك على الفور قيداً مصطنعاً غير مرغوب فيه على المستفيد. وفي ظل هذه الظروف تدخل الحاجة إلى المعلومات قسراً، وربما دون وعي، في لغة النظام، وربما قبل المستفيد بما هو أقل دقة أو أقل اكتمالاً مما يريده فعلاً.

ومن الممكن توضيح هذه الظاهرة بمثال بسيط جداً؛ فمن الممكن لمستفيد أن يلجأ إلى أحد مراكز المعلومات ويعلن عن اهتمامه بتصنيع أنابيب الصلب غير القابل للصدأ بعمليات اللحام. ويراجع المستفيد اختصاصي المعلومات لغة النظام معاً، مع الاهتمام بوجه خاص بمصطلحات المعادن ومصطلحات اللحام ومصطلحات المنتجات المصنعة. ونظراً لأن المستفيد يجد مصطلح الحديد غير القابل للصدأ STAINLESS STEEL (بلا مزيد من التفريع) في اللغة فإنه يقبله إلى جانب مصطلح الأنابيب TUBES. وبمراجعة مصطلحات اللحام يجد أن المصطلح المخصص «لحام القوس المدرع SHIELDED ARC WELDING» قد ورد كتفريع للحام، فيقبل المصطلح المخصص. فالطلب الذي يسلم لمركز المعلومات إذن هو البحث عن تصنيع أنابيب الصلب غير القابل للصدأ بعمليات لحام القوس المدرع. وللأسف فإن هذا البحث لن يكون مرضياً تماماً؛ فالمستفيد يهتم فعلاً بشيء أكثر تخصيصاً، وهو المعلومات حول نوع من الصلب غير القابل للصدأ بتركيبية معينة (الكروم - النيكل - الفاناديوم) وطريقة اللحام المحددة المسماة بلحام القوس بالهليوم. والواقع أنه بإمكان النظام البحث عن هذا الموضوع بشكل مخصص باستعمال مصطلحات السبائك الخاصة بالكروم والنيكل والفاناديوم، وبإضافة مصطلح الغاز هليوم HELIUM، إلا أن المستفيد لا يدرك ذلك. ومن ثم يبدأ إجراء بحث أكثر عمومية من الحاجة، فقد تأثر المستفيد هنا سلباً بتفسيره لإمكانات النظام.

وينبغي أن يكون من القواعد التي لا محيد عنها أن يطلب من المستفيدين من مراكز المعلومات كتابة بيان بحاجتهم إلى المعلومات بلغتهم، وألا يحاولوا، في المراحل الأولية على الأقل، التعبير عن حاجتهم بمصطلحات النظام. وهذا الأمر مهم لسبب آخر خلاف قضية وضوح الطلب. فإذا ما طلب من المستفيدين التعبير عن احتياجاتهم بمصطلحات مختارة من لغة النظام المقيدة، فلن يكون من الممكن أبداً التحقق من المجالات التي تفتقد فيها اللغة التخصيص الكافي. فإذا تقيّد المستفيدون مثلاً بالسؤال عن لحام القوس

المدرع بينما هم يريدون فعلاً البحث بشكل أكثر تخصيصاً عن لحام القوس بالهليوم أو لحام القوس بالأرجون ، فإنه لن يتم إدراك الحاجة إلى إضافة هذه المصطلحات الأكثر تخصيصاً إلى اللغة . ويمكن في النهاية لمركز المعلومات أن يفقد كثيراً من المتعاملين معه لأن لغته لم تحرص على درجة كافية من التخصيص تكفل مساهمة التطورات الحديثة في المجالات التي تغطيها ، ومن ثم لا يمكن إجراء عمليات البحث بدرجة كافية من الدقة ترضي المستفيدين .

وربما تعطي هذه المناقشة الانطباع بأنه من الممكن لأي شكل من أشكال التفاعل بين المستفيد واختصاصي المعلومات ، بهدف تحسين أو توضيح بيان الحاجة ، أن يؤدي إلى انخفاض مستوى هذا البيان ، والابتعاد به كثيراً عن الطلب الفعلي . وليس الأمر كذلك؛ وإنما يبدو أن هناك مرحلة لا ينبغي تجاوزها ، يمكن أن يحدث عندها هذا التفاعل.

ونظراً لما انتهى إليه من نتائج في تقييم المدلرز ، أكد لانكستر (1968 a) **Lancaster** أهمية نموذج طلب البحث الذي يتم تصميمه جيداً . وهو يفترض أنه من الممكن لتتابع التفاعلات ، من زيارة المستفيد لمركز المعلومات ، وأن يطلب من المستفيد تعبئة نموذج الطلب ، ويجري الوسيط مقابلة مع المستفيد لاستيضاح وتنقية ما سجل في النموذج ، أن يسفر بوجه عام عن نتائج أفضل من التتابع التالي : زيارة المستفيد للمركز ، وإجراء الوسيط للمقابلة مع المستفيد ، وتعبئة نموذج الطلب بواسطة الوسيط أثناء إجراء المقابلة.

وقد ظل هذا الفرض دون اختبار إلى أن قارن فترزجيرالد (1981) **Fitzgerald** بين التسلسلين في التفاعل ، في إحدى المكتبات الطبية . فاعتماداً على وسيط واحد ، قام بإجراء المقابلات مع المستفيدين وإجراء عمليات البحث أيضاً ، كان في متناول فترزجيرالد مجموعة من الطلبات التي عوملت وفقاً للطريقة الأولى ومجموعة ثانية عوملت بالطريقة الأخرى ، واتبع أحد سبل التوزيع العشوائي في تقسيم الطلبات على المجموعتين . ثم قام بعد ذلك بمقارنة نتائج عمليات البحث بناء على الاستدعاء والتحقيق وتكلفة البحث ، وتكلفة الوحدة لكل وثيقة متصلة بالموضوع تم استرجاعها . ولم تؤيد هذه البيانات ، الفرض ، نظراً لأنه بالنسبة لمعظم معايير التقييم ، لم تكن

النتائج التي تحققت في أي من طريقتي التفاعل ، تختلف اختلافاً جوهرياً عن تلك التي تحققت في الأخرى . إلا أنه قد تصادف وقوع حدث مهم في هذه المكتبة الطبية ، لم يكن الباحث على علم به أثناء التخطيط للدراسة ؛ حيث كان المستفيد ومسئول البحث يجلسان دائماً معاً إلى المنفذ أثناء إجراء عملية البحث ، وبذلك كان البحث إيعازياً فعلاً ، ينطوي على تفاعل بين المستفيد ومسئول البحث ومرصد البيانات . ومن الممكن أن يكون هذا شكلاً أكثر فعالية بكثير من أشكال التفاعل الأخرى التي محت تقريباً أي وجه للاختلاف كان من الممكن أن يحدث في تفاعل ما قبل إجراء البحث .

والموقف الذي يجلس فيه كل من المستفيد والوسيط معاً أمام المنفذ أثناء إجراء البحث هو المفضل بشكل ظاهر . وما لم يكن ذلك ممكناً فإنه من الممكن إجراء شكل ما من أشكال البحث المتزامن عن بعد (Graham, 1980) وفي هذه الحالة هناك فاصل جغرافي بين المستفيد ومسئول البحث ، إلا أن الهاتف يتيح لهما فرصة الاتصال ، كما تتيح الارتباطات الشبكية للمستفيد القدرة على أن يلاحظ عن طريق منفذه ما يقوم به مسئول البحث وما يتم استرجاعه . وإذا لم يكن من الممكن توفير أي شكل من أشكال التفاعل بين المستفيد والوسيط أثناء إجراء البحث ، فإنه يمكن حينئذ أن يكون التفاعل المفضل قبل إجراء البحث ، هو التفاعل الذي يتم فيه تعبئة نموذج طلب البحث من جانب المستفيد قبل حدوث أي مناقشة بين المستفيد والوسيط .

نماذج طلب البحث :

من الوظائف المهمة لمرفق المعلومات مساعدة المستفيدين من خدماته في صياغة الطلبات الملائمة ، ويمكن لهذه المساعدة أن تتخذ عدداً من الأشكال ، بدءاً بالإرشادات العامة حول خصائص النظام وإمكاناته ، من خلال الموجز الإرشادي للمستفيدين مثلاً ، إلى إعداد نموذج لطلب البحث مصمم لمساعدة المستفيد في التعبير عن حاجته بدقة . ويمكن لنموذج طلب البحث ، إذا أحسن تصميمه أن يفيد في تجميع البيانات التي يمكن أن يعتمد عليها اختصاصيو المعلومات في تفسير احتياجات المستفيدين . فضلاً عن البيان الكامل للطلب باللغة الطبيعية ، يمكن لمثل هذا النموذج أن يشتمل على بيانات حول الهدف من البحث ونوعية البحث المطلوب (شامل أو انتقائي) . كما يمكن أيضاً أن يمد المستفيد بسلسلة من خانات المراجعة التي يمكن بواسطتها تحديد مجال البحث

بالشكل المناسب . فمن الممكن في مجال الطب على سبيل المثال ، استعمال خانات المراجعة في بيان ما إذا كان المستفيد يهتم فقط بالدراسات المتصلة بالإنسان ، أم بالدراسات التي تتصل بالحيوان أم بكليهما ، وما إذا كان المستفيد يهتم فقط بجنس أو نوع بعينه أو بفئة عمرية معينة ، وما إذا كان يرغب في دراسات الحالة ، أم في الدراسات التي تتناول الأحياء أم الدراسات المعتمدة على التشريح بعد الوفاة . ويمكن للنموذج الذي أحسن تصميمه أن يساعد المستفيد في التعبير عن طلبه بشكل صريح مكتمل . كما يمكن أن يساعد المستفيد بوجه خاص في استبعاد مختلف المجالات أو الجوانب التي لا تهمه ، كالدراسات الخاصة بالحيوانات مثلاً . وربما لا يتنبه المستفيد لمثل حالات الاستبعاد هذه دون مساعدة النموذج . كذلك ينبغي أن يشتمل النموذج على حيز يسمح للمستفيد بتسجيل الإشارات الوراقية الخاصة بالوثائق التي يعرف فعلاً أنها تتصل بموضوع طلبه ، إذا ما توافرت هذه الإشارات . ويمكن لهذه الإشارات أن تخدم أكثر من هدف واحد ؛ حيث تفيد في التأكد ما إذا كانت صيغة البحث في المستوى المناسب من التخصيص ، وفي التأكد من مدى اكتمال البحث (هل استرجعت الوثائق المعروفة المتصلة بموضوع الطلب ، على افتراض وجودها في مرصد البيانات ؟) كما يمكن أن تتخذ أساساً لاستراتيجية البحث (استرجاع تسجيلات الوثائق التي نعرف أنها متصلة بموضوع الطلب وصياغة استراتيجية للبحث عن التسجيلات الأخرى المماثلة لها) . ونتيجة لتقييم المدرز تم تصميم نموذج طلب بحث يشتمل على جميع العناصر التي عرضنا لها ، وذلك للمكتبة القومية للطب (Lancaster, 1971) . وكان لهذا النموذج أثره في كثير مما تلاه . ويوضح الشكل رقم (٤٥) الذي يرد فيما بعد في هذا الكتاب ، نموذجاً تتوافر فيه جميع الخصائص المطلوبة .

الفصل الخامس

التحليل الموضوعي

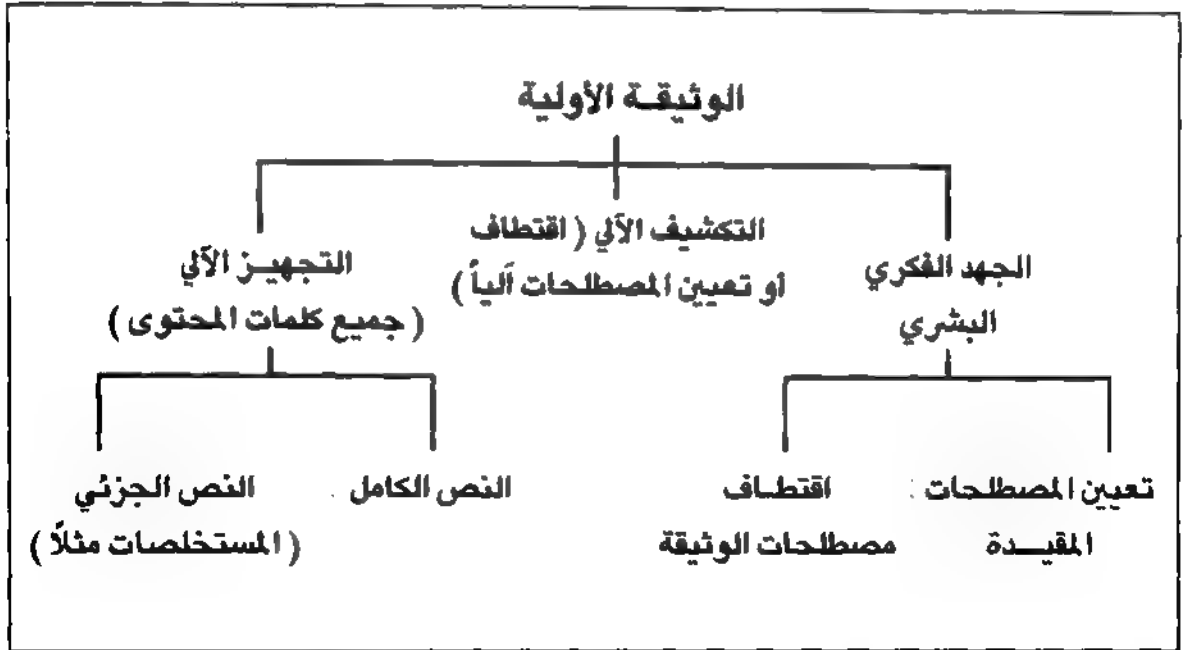
تحتل طرق وأشكال التعبير عن موضوعات الوثائق بؤرة الاهتمام في تصميم نظم استرجاع المعلومات (انظر الشكلين رقمي (٣) و (١٦)). وللقرارات الخاصة بالتحليل الموضوعي والتعبير عن ناتج التحليل أهميتها البالغة ، نظراً لأنها تؤثر في عمليات البحث التي يمكن إجراؤها في النظام . وعلى الرغم من أن توفير المعلومات الموضوعية يمثل بؤرة الاهتمام في صناعة استرجاع المعلومات ، فإن المدخل الموضوعي في فهرس المكتبات لم يكن دائماً يحظى بما يستحق من اهتمام . ومن يمن الطالع ، أن مجال المكتبات قد بدأ يكرس المزيد من الاهتمام لمشكلة المدخل الموضوعي منذ بدأت الفهارس المتاحة للجمهور على الخط المباشر تحل محل الفهارس البطاقية .

ويقدم الشكل رقم (٢٤) البدائل المحتملة للتعبير عن ناتج التحليل الموضوعي . ودائماً ما يبدأ التحليل الموضوعي والتعبير عن ناتج التحليل بوثيقة أولية ، إلا أنه يتبين من الشكل أن مختلف طرق المعالجة الموضوعية تختلف فيما بينها اختلافاً جوهرياً . ومن الممكن تلخيص أوجه الاختلاف هذه على النحو التالي :

١ - يمكن للتحليل الموضوعي أن يتم بالجهد الفكري البشري أو بواسطة بعض خوارزميات الحاسب الآلي .

٢ - يمكن للتعبير عن الموضوعات أن يعتمد على لغة الوثيقة الأولية ، أو على بعض اللغات الأخرى .

٣ - يمكن للتعبير عن الموضوعات أن يكون ناتج اختيار جميع الكلمات (تصويرياً) الواردة في النص الكامل أو الجزئي للوثيقة ، أو الاختيار من بين الكلمات أو المفاهيم أو الموضوعات الواردة في الوثيقة ، وفق أسس معينة .



الشكل رقم (٢٤) : مختلف الطرق المحتملة للتعبير عن الموضوع .

وهكذا ، يمكن لمصطلح « التكشيف » أن يدل على :

١ - عملية تحليلية تنطوي على تعرف البشر على موضوعات الوثيقة ، والاختيار من بين هذه الموضوعات ثم التعبير عن الناتج (Pao, 1989) . ويمكن لذلك أن ينطوي على ترجمة المفاهيم أو الموضوعات التي تم التحقق منها إلى مصطلحات لغة مقيدة ، أو اقتطاف المصطلحات من الوثيقة الأولية .

٢ - عملية آلية يتم فيها اقتطاف مجموعة فرعية من المصطلحات التي تعد ممثلة للمحتوى الموضوعي للوثيقة .

٣ - عملية آلية يتم فيها وضع جميع كلمات المحتوى (*) في جميع المجالات المتصلة بالموضوع (المصطلحات الكشفية ، أو العنوان ، أو المستخلص ، أو النص الكامل) في ملف مصنف لأغراض البحث على الخط المباشر .

(*) ما يدل على المحترى . (المترجم) .

ويتناول هذا الفصل القضايا والتعريفات والسياسات والمشكلات المرتبطة بالنوعين الأول والثالث للتكشيف، أما النوع الثاني فيمثل بؤرة الاهتمام في الفصل الحادي عشر .

التكشيف بواسطة البشر :

يمكن للتكشيف أن ينطوي على التقاط الكلمات والعبارات من الوثائق مباشرة (التكشيف الاشتقاقي) أو التعبير عن المحتوى الموضوعي بمصطلحات مستقاة من لغة مقيدة (التكشيف بالتعيين) الشكل رقم (٢٤) . وفي التكشيف بالتعيين ينطوي التحليل الموضوعي والتعبير عن ناتج التحليل على عدة خطوات ، يتخذها المكشف أو المفهرس ، تبدأ بتحليل الوثيقة للتعرف على المفاهيم أو الموضوعات القابلة للتكشيف (التحليل الموضوعي أو تحليل المفاهيم) ، ثم تحول هذه الموضوعات إلى لغة تكشيف مناسبة (الترجمة) . ويسفر ذلك عن مجموعة من المصطلحات قصد بها وصف المحتوى الموضوعي للوثيقة ، والتي يمكن أن تؤدي إلى استرجاعها عندما تضاهى مقابل المصطلحات نفسها في الاستفسار أو استراتيجية البحث . و وراء هذه الخطوات عدة سياسات توجه عملية التكشيف ونتيجتها النهائية والتي غالباً ما تؤثر في كفاءة نظام الاسترجاع .

ومن بين السياسات ما يحدد المادة القابلة للتكشيف في الوثيقة . وعادة ما يكون المجال القابل للتكشيف بالنسبة للوثائق التي تغطيها معظم مراصد البيانات الوراقية، هو نص الوثيقة كاملاً . ومن ثم فإنه يتعين على المحلل أن يبنّي قرارات التكشيف على الوثيقة ككل لا على قطاعات منها فقط ، كالعنوان أو المستخلص . هذا بالإضافة إلى أن المحللين غالباً ما يُطالبون بتكشيف شكل المادة أيضاً ، أي (في مقابل ما تتناوله أو تدور حوله) . فإذا كانت الوثيقة محل النظر ، على سبيل المثال ، وراقية لمصادر المعلومات حول موضوع الأمية ، فإنها يمكن أن يعين لها مصطلحات تدل على كل من شكلها (وراقية) وموضوعها (الأمية) .

وهناك عدة جوانب فكرية في الوثيقة يمكن أن تعد قابلة للتكشيف . فغالباً ما تنص سياسات التكشيف مثلاً على ضرورة التحقق من بعض العناصر كالنوع (مرصد بيانات الأدب) أو الفترة الزمنية (مرصد بيانات التاريخ) أو المجموعة المستهدفة بالتعليم (مرصد بيانات التعليم) التحقق دائماً من هذه العناصر وتعيين المصطلحات

الكشفية المناسبة لها . إلا أن الجهد الرئيس في المعالجة الموضوعية ينطوي على التحقق من الموضوعات التي تتناولها الوثيقة ، أي ما تدور حوله الوثيقة ، والتعبير عن هذه الموضوعات . وتؤكد السياسات في هذا المجال المهم على أن المحتوى الموضوعي للوثيقة هو ما ينبغي تحليله والتعبير عنه بمجموعة من المصطلحات الكشفية . ويسمى هاتشنز (1978) Hutchins هذا بالمدخل التلخيصي للتحليل الموضوعي .

وتهتم سياسة أخرى بالشمول أو الإحاطة exhaustivity في الكشف ، والشمول هو مدى الإحاطة بكل الموضوعات المتميزة التي تحظى بالمعالجة في وثيقة معينة ، في عملية الكشف ، وترجمة هذه الموضوعات إلى لغة النظام . ونتناول ما لهذه السياسة من تأثير على كفاءة نظام الاسترجاع ، بمزيد من التفصيل ، في الفصل العاشر . ويميل الكشف المفرق في الشمول عموماً لارتفاع بمستوى الاستدعاء ، إلا أنه يؤدي إلى انخفاض التحقيق ، بينما يؤدي الكشف المفرق في الانتقال إلى الارتفاع بمستوى التحقيق ، إلا أنه لا يكفل الاستدعاء المرتفع . وفي الكشف الشامل يتم تعيين عدد كبير من المصطلحات الكشفية لكل وثيقة ، لأن المكشفين مطالبون بتعيين مصطلحات لكل من الموضوعات الرئيسة والموضوعات الفرعية للوثيقة ، أي الموضوعات البؤرية بالنسبة للوثيقة (الرئيسة) والموضوعات التي تعالج بإيجاز فقط (الفرعية) . وهكذا ، فإن مسئول البحث عادة ما يسترجعون مجموعات كبيرة من الوثائق التي تشتمل على وثائق تتناول الموضوع كقضية بؤرية أو مركزية ، بالإضافة إلى الوثائق التي تتناوله بإيجاز (استدعاء مرتفع) . أما في الكشف الأقل شمولاً (الأكثر انتقائية) فعلى العكس ، يميل مسئولو البحث إلى استرجاع مجموعات صغيرة من الوثائق التي لا تضم سوى الوثائق التي تتناول الموضوع كقضية مركزية (تحقيق مرتفع) .

ودرجة الشمول أحد وجهي الاختلاف الرئيسين بين ما يسمى عادة « الفهرسة الموضوعية » وما يسمى « الكشف الموضوعي » ، أما وجه الاختلاف الآخر فهو درجة الربط في لغة الكشف (انظر الفصل السادس) . ويعرض الشكل رقم (٢٥) تسجيلية وراقية من فهرس مكتبة وأخرى من مرصد بيانات إريك ERIC . ويتضح لأول وهلة أن الشمول في التسجيلية الثانية أعلى مما هو عليه في الأولى ، حيث تميل الفهارس لتقديم مداخل موضوعية مهمة جداً للوثائق الكاملة (كالكتب الكاملة أو الدوريات أو أعمال

Meadow, Charles T.

Basics of online searching / Charles T. Meadow, Pauline (Atherton) Cochrane.

-- New York : Wiley, 1981.

xiv, 245 p. : ill. ; 26 cm. -- (Information sciences series)

"A Wiley-Interscience publication."

Includes bibliographical references and index.

ISBN 0-471-05283-3

SUBJECT HEADINGS (Library of Congress; use s=):

On-line bibliographic searching

MIRLYN Online Catalog, University of Michigan

ED205181# IR009484

Basics of Online Searching.

Meadow, Charles T.; Cochrane, Pauline (Atherton)

Wiley (John) & Sons, New York, N.Y.

1981

255p.; A Wiley-InterScience Publication.

Report No.: ISBN-0-471-05283-3

Available from: Wiley-InterScience, 605 Third Ave., New York, NY 10158.

Document Not Available from EDRS.

Language: English

Document Type: BOOK (010); CLASSROOM MATERIAL (050)

Geographic Source: U.S.; New York

Journal Announcement: RIEDEC81

Intended to teach the principles of interactive bibliographic searching to those with little or no prior experience, this textbook explains the basic elements of online information retrieval and compares the major database search systems. Its chapters address (1) relevant definitions and vocabulary; (2) the conceptual facets of database searching, search formulation, and online costs; (3) the presearch interview; (4) terminals and networks; (5) search languages; (6) database organization and record structures; (7) basic system commands; (8) text searching; (9) beginning and ending a search; (10) storing searches and selective dissemination of information (SDI); (11) search aids; and (12) search strategies. Appendices provide a summary of search languages, a list of available online databases, and examples of database descriptions and search aids. Examples are based on the BRS, DIALOG, and ORBIT search systems. (FM)

Descriptors: Databases; *Information Retrieval; Information Systems; *Online Systems; *Search Strategies

Identifiers: Command Language; Search Keys; Search Negotiation (Computer Science)

ERIC database (*Current Index to Journals in Education*)

الشكل رقم (٢٥) : الشمول في التكشيف في تسجيلية فهرس مقارناً بالشمول في مرصد بيانات إرك .

المؤتمرات مثلاً) ، بينما يمكن للكشافات أن تشير إلى الوثائق الكاملة أو إلى أجزاء الوثائق (كالمقالات أو البحوث أو الفصول مثلاً) التي يتم تحليلها بشكل أكثر شمولاً (إحدى عشرة واصفة تقريباً للوثيقة في إريك ERIC) مما هو مألوف بالنسبة لفهارس المكتبات (حيث يتراوح المتوسط بين رأس موضوع واحد وثلاثة رؤوس للتسجيلية) .

وفي تسجيلية إريك استعمل المكشف ثمانية مصطلحات كشفية (« واصفات » و « محددات ») كما ميز تلك التي تتمتع بأهمية خاصة بنجمة . وإذا نظرنا إلى الوثيقة التي أعطيت واصفات أساسية وأخرى فرعية باعتبارها أقصى درجات التكشيف شمولاً ، فإنه يمكن التعرف على معالم ثلاثة مستويات للشمول في الشكل رقم (٢٥) :

١ - المحتوى الكلي للوثيقة فقط ، ومن ثم أقل درجات الشمول (تسجيلية فهرس ميرلين (MIRLYN) .

٢ - الواصفات الرئيسية فقط (الواصفات المميزة بنجمة في تسجيلية إريك) .

٣ - كل من الواصفات الرئيسية والفرعية، ومن ثم الأكثر شمولاً (تسجيلية إريك كاملة) .

وهكذا يميل المحلل الذي يقوم بالتكشيف لصالح منتج مرصد البيانات للتعبير عن ناتج التعرف على المحتوى الموضوعي للوثائق بتفصيل أكثر من المحلل (الم فهرس) الذي يضطلع بالمهمة المناظرة في فهرس المكتبة . وكان المستوى المنخفض للشمول يعد في الماضي مناسباً تماماً للفهارس ، التي ظلت لسنوات عديدة تقتصر على الشكل الورقي الجامد المعتمد على الربط المسبق ، حيث لم تكن تتاح للمستفيدين إمكانية الربط بين رؤوس الموضوعات أثناء البحث عن موضوعات أكثر تخصيصاً من تلك التي تغطيها رؤوس الموضوعات المفردة .

وعلى الرغم من أن الجمع بين الواصفات الرئيسية والفرعية يمثل أقصى درجات الشمول في التكشيف بالنسبة لهذه الوثيقة ، فإنه من المفيد التمييز بين الموضوعات الرئيسية والموضوعات الفرعية : فلا تحظى جميع المفاهيم أو الموضوعات الواردة في الوثيقة بالتغطية بالقدر نفسه من التفصيل ، ومن ثم فإنها لا تستحق جميعها المعاملة نفسها . ويكفل التمييز الصريح للواصفات الرئيسية في بعض مراصد البيانات لمسئولي البحث

فرصة انتقاء الوثائق التي تتصل بشكل رئيس بموضوع الطلب (تحقيق مرتفع)، أو تتصل بشكل رئيس وفرعي في الوقت نفسه (استدعاء مرتفع). والتمييز بين الواصفات الرئيسية والواصفات الفرعية إنما هو في الواقع محاولة أولية « للتكشيف الموزون » .

ومن المصطلحات المألوفة في الإنتاج الفكري ، إلا أننا تجنبناه عمداً في هذا الكتاب «العمق» . وكما يستعمل في الإنتاج الفكري ، فإن التكشيف المتعمق يعني مجرد استعمال عدد من المصطلحات أكبر مما يستعمل في التكشيف غير المتعمق . وسواء كانت المصطلحات الإضافية تستعمل لتغطية المزيد من الموضوعات (زيادة الشمول) أو لتكشيف عدد محدود من الموضوعات بطريقة أكثر تحديداً (زيادة التخصيص) أمر قلما ننتبئه بوضوح . وبعبارة أخرى ، فإن التكشيف المتعمق يُستعمل من جانب بعض المؤلفين لوصف مدى الدقة في تحديد الفئات ، بينما يُستعمل من جانب آخرين لوصف مدى الإحاطة أو التعرف على جميع الموضوعات التي تتناولها الوثيقة عند تكشيفها.

وتهتم إحدى سياسات التكشيف الأخرى بالتخصيص ؛ فعلى المكشف أو المفهرس أن يعين لكل موضوع قابل للتكشيف المصطلح أو المصطلحات التي تتساوى مع الموضوع في مستوى التخصيص . وعلى ذلك ، فإنه إذا كانت هناك وثيقة عن فهارس الخط المباشر ، فإنه ينبغي استعمال المصطلح « فهارس الخط المباشر ONLINE CATALOGS لا المصطلح « فهارس CATALOGS » . وتيسر هذه السياسة عمليات البحث الدقيقة على اختلاف مستويات التخصيص ، أي أنه من المفترض أن يستعمل المستفيدون الذين يريدون معلومات عن فهارس الخط المباشر ذلك المصطلح ، أما إذا رغبوا في توسعة البحث (زيادة الاستدعاء) فإنهم يضيفون إلى استراتيجية البحث الخاصة بهم المزيد من المصطلحات من الفئة الموضوعية نفسها والواردة في اللغة المقيدة (أعرض أو أضيق أو في المستوى نفسه في التقسيم الهرمي) .

وتهتم إحدى القضايا المتصلة بسياسة التكشيف بطبيعة اللغة المستخدمة في وصف المحتوى الموضوعي للوثائق ؛ ويمكن لهذه اللغة أن تكون مقيدة ، أو شبه مقيدة ، أو غير مقيدة . ويسمى التكشيف الذي يستخدم اللغة المقيدة (كالمكنز أو قائمة رءوس الموضوعات) بالتكشيف بالتعيين . ويستخدم التكشيف بالتعيين من جانب العديد من مؤسسات التكشيف والاستخلاص بالإضافة إلى المكتبات . ولا يعتمد التكشيف في

بعض مراصد البيانات الوراقية على اللغات المقيدة ، وإنما ينطوي على انتقاء المصطلحات الواردة في الوثيقة نفسها . ويسمى هذا بالتكشيف بالاقتطاف أو التكشيف الاشتقاقي .

والافتقار التام إلى اللغة المقيدة في التكشيف بواسطة البشر نادر ، لأن المكشفين عادة ما يكونون مطالبين بأن يستعملوا على الأقل أوزاناً صرفية معينة (أسماء في العادة) للمصطلحات التي يختارونها . ومن الممكن ممارسة المزيد من التقييد بإقرار المزيد من القواعد المحكمة لصياغة المصطلحات في إطار مختلف الفئات ، كأسماء الأعلام من الأفراد والمؤسسات . ووجه الاختلاف الجوهرى بين نظم التكشيف المعتمدة على اللغات المقيدة والنظم المعتمدة على اللغات شبه المقيدة وتلك المعتمدة على اللغات غير المقيدة ، هو أن اللغة المقيدة تتكون من قائمة نهائية منشورة ، أما اللغة شبه المقيدة أو اللغة غير المقيدة فيمكن أن تسمح بإضافة المصطلحات الجديدة بمجرد ظهورها طالما كانت مفيدة في إطار الوثائق التي يتم تكشيفها . ويتبين من تسجيلة إرك الواردة في الشكل رقم (٢٥) وجود حقل (للمحددات) شبه المقيدة ، يشتمل على نماذج لهذا النوع من المصطلحات . وقد كشفت الوثائق في هذا الملف بلغة مقيدة وأخرى شبه مقيدة .

ولما كان التكشيف بالجهد الفكري البشرى مرتفع التكلفة ويستغرق وقتاً طويلاً ، فقد تم توجيه قدر كبير من جهود البحث للتأكد من مدى فعاليته . وقد أمكن التحقق من عدد من القضايا المهمة ومناقشتها في الإنتاج الفكري . وقد أثار بعض المؤلفين التساؤل حول الافتراض الأساس الذي يقوم عليه التكشيف ، وهو أن الوثائق «تدور حول» موضوع ما أو عدد من الموضوعات ، وأن مهمة المكشف تلخيص هذه الموضوعات في مجموعة من المصطلحات المناسبة . إلا أن كثيراً من الباحثين ، ومن بينهم مارون (1977) Maron وسويقت وون وبريمر (Swift, Winn and Bramer 1978) وهاتشنز (1978) Hutchins أشاروا إلى المشكلات المرتبطة بهذا الافتراض ، مدعين أن خصائص المستفيدين وسلوكهم في البحث عن المعلومات تشكل معايير أكثر أهمية مما تدور حوله الوثائق ، في تحديد أي الوثائق يمكن أن تلبي حاجة معينة إلى المعلومات . وتدل مقترحاتهم على أننا ينبغي أن نكشف من أجل المستفيدين المزمع خدمتهم لا من أجل الوثيقة . وقد حاول منتج مراصد البيانات ذلك على نطاق محدود ، إلا أن تلخيص الوثائق بوجه عام ما زال هو الأساس المحوري في التكشيف .

التكشيف الجزئي وتكشيف النصوص كاملة :

في الطرف الآخر لطيف التكشيف نجد الإجراءات الآلية المطلقة التي توضع فيها جميع الكلمات الدالة على المحتوى ، الواردة في الحقول المتصلة بالموضوع ، في ملف مصنف لأغراض البحث على الخط المباشر . وفي هذه الحالة يستخدم متعهد مراصد البيانات ما يعرف بخوارزميات الإعراب **parsing algorithm** لتجهيز البدائل الوراقية أو النصوص الكاملة . وعلى الرغم من اختلاف هذه الخوارزميات في دقائقها من متعهد إلى آخر ، فإنها دائماً ما تتوافر بها الخصائص التالية :

- ١ - أنها تختار فقط المصطلحات ذات الدلالة من التسجيلية باستخدام قائمة استبعاد **stop list** (تتكون أساساً من الكلمات الوظيفية كالأدوات وحروف الجر والظروف) لاستبعاد المصطلحات غير المرغوب فيها .
 - ٢ - أنها تجهز الحقول المتصلة بالموضوع ، بما في ذلك العناوين ، والواصفات ، والمحددات **identifiers** إن وجدت ، والمستخلصات ، والنصوص الكاملة .
 - ٣ - أنها تضع هذه المصطلحات الدالة على المحتوى في كشاف مصنف ، إلى جانب المعلومات الخاصة بالحقول التي التقطت منها ومواقعها المحددة في تلك الحقول .
- ويوضح المثال الوارد في الشكل رقم (٢٦) كيف يتم تنفيذ هذه العملية ، بتقديم تسجيلية وراقية من إرك وما يرتبط بها من مداخل الكشاف المصنف **inverted** في نظام ديالوج **DIALOG** . ويتم التقاط المصطلحات المفردة التي ترد في الكشاف المصنف من العناوين والمستخلصات والمحددات والواصفات . أما العبارات التي ترد في هذا الكشاف فهي واصفات أو محددات كاملة . ويمكن لمسئولي البحث انتقاء كلمات مفردة أو عبارات من أي حقل أو أية حقول في تسجيلية مرصد البيانات ، وتسمى مضاهاة أي مصطلح أو عبارة في أي من هذه الحقول بالبحث في النصوص المطلقة . ويأتى ذلك في مقابل عمليات البحث التي تتطلب المضاهاة الدقيقة لوصفة أو محدد بأكمله ، وغالباً ما يسمى هذا بالبحث باللغة المقيدة .

ويمكن القول إذن في الواقع أن الوثيقة تتناول تلك الموضوعات التي تم التعبير عنها بمجموع الكلمات والعبارات ذات الدلالة الواردة في كل هذه الحقول . ومن الممكن

Bibliographic Record (from CUE, 1979)

EJ206829 IR507016

Title: Status of Stasis: Academic Librarians 10 Years Later.

T11 T12 T13 T14 T15 T16 T17 T18

Author: Galloway, R. Dean

Source: American Libraries, v10 n6 p349-352 Jun 1979

Available from: Reprint, UMI

Language: English

Document Type: JOURNAL ARTICLE (080): HISTORICAL MATERIAL (060):

POSITION PAPER (120)

Journal Announcement: CUDDEC79

Abstract: Despite higher qualifications, academic librarians still

AB1 AB2 AB3 AB4 AB5 AB6

lack equality with faculty.

AB7 AB8 AB9 AB10

Descriptors: *Academic libraries; Attitude change; Historical

DE1 DE2 DE3 DE4 DE5

reviews: Librarians; Professional continuing education;

DE6 DE7 DE7 DE8 DE9 DE10

Professional recognition; *Qualifications; Research; *Status

DE11 DE12 DE13 DE14 DE15

Additions to inverted index:

| | | |
|-----------------------------------|----------|------------|
| 10 | EJ206829 | T16 |
| academic | EJ206829 | T14 |
| | EJ206829 | AB4 |
| academic libraries | EJ206829 | DE1 |
| attitude | EJ206829 | DE1DE2 |
| attitude change | EJ206829 | DE3 |
| change | EJ206829 | DE3DE4 |
| continuing | EJ206829 | DE4 |
| despite | EJ206829 | DE9 |
| education | EJ206829 | AB1 |
| equality | EJ206829 | DE10 |
| faculty | EJ206829 | AB8 |
| higher | EJ206829 | AB10 |
| historical | EJ206829 | AB2 |
| | EJ206829 | DE5 |
| historical review | EJ206829 | DE5DE6 |
| lack | EJ206829 | AB7 |
| later | EJ206829 | T18 |
| librarians | EJ206829 | T15 |
| | EJ206829 | AB5 |
| libraries | EJ206829 | DE7 |
| professional | EJ206829 | DE2 |
| | EJ206829 | DE8 |
| professional continuing education | EJ206829 | DE11 |
| professional recognition | EJ206829 | DE8DE9DE10 |
| qualifications | EJ206829 | DE11DE12 |
| | EJ206829 | AB3 |
| recognition | EJ206829 | DE13 |
| research | EJ206829 | DE12 |
| reviews | EJ206829 | DE14 |
| stasis | EJ206829 | DE6 |
| status | EJ206829 | T13 |
| | EJ206829 | T11 |
| still | EJ206829 | DE15 |
| with | EJ206829 | AB6 |
| years | EJ206829 | AB9 |
| | | T17 |

تلخيص مزايا هذه الطريقة وعيوبها بمقارنة هذه العملية الآلية بالعملية الفكرية البشرية التي سبق أن تناولناها . ومن بين نتائج التكشيف بواسطة البشر أنه من الممكن غالباً التعبير عن فكرة أو موضوع معين بمصطلح واحد فقط ، قد يكون هو المصطلح نفسه الذي يستعمله مسئول البحث ، وقد لا يكون كذلك . ويمكن لمعالجة حقول مرصد البيانات على هذا النحو أن تساعد في استرجاع الوثائق المتصلة بالموضوع ، نظراً لأنه من الممكن ، على سبيل المثال ، لمسئول البحث أن يستعمل مصطلحاً يضاهي كلمة أو عبارة في المستخلص ، ولا يضاهي الوصفة التي يستعملها المكشف . ويدعم هذا ، الرأي القائل بأن عمليات تجهيز واختزان واسترجاع التسجيلات الوراقية ، تعتمد على الاعتقاد بأن الحشو اللغوي من السمات المرغوبة في استرجاع المعلومات ، حيث يمكن للتعبير عن فكرة معينة بأكثر من شكل لغوي واحد في المجالات المختلفة أن ييسر استرجاع الوثائق المتصلة بالموضوع . ويرتبط بهذا الرأي ارتباطاً وثيقاً أن اللغة المستعملة في حقول كالعنوان والمستخلص والنص الكامل أحياناً ما تكون مخصصة إلى حد بعيد ، ومن ثم فإنها يمكن أن تكون مفيدة أكثر من غيرها بالنسبة لمسئولي البحث الذين يتلقون طلبات تتعلق بمعلومات مخصصة جداً .

كذلك يمكن مقارنة هذه الطريقة بالتكشيف البشري بناء على كل من الوقت والتكلفة . فوجود عنصر بشري في التكشيف يزيد وبشكل ملحوظ ما يتحملة المنتج من تكلفة مرصد البيانات ، لأنه يتطلب الاستعانة باختصاصيين ، فضلاً عن الموارد المادية اللازمة للمحافظة على تجدد اللغة المقيدة ، واستخدام هذه اللغة في تحليل الوثائق الواردة . أما بالنسبة للمنتجين الذين ينشئون مراصد البيانات دون الحاجة إلى تسجيلات ناتجة عن التكشيف بواسطة البشر ، فإن غياب هذا العنصر المستنزف للموارد عادة ما يسفر عن اقتصاد ملحوظ في التكلفة . وهناك أيضاً بعض الحلول التوفيقية بالنسبة لمسئول البحث : فعادة ما يتطلب البحث في مراصد البيانات التي لا تستعمل مصطلحات مقيدة المزيد من الوقت والجهد ، نظراً لأنه سيكون على مسئول البحث أن يتدبر ، وأن يدخل في استراتيجيات البحث جميع المترادفات الخاصة بفكرة أو مفهوم معين ، أي أن التحكم في الترادف يأتي في مرحلة البحث . وفي مقابل ذلك ، فإن مسئول البحث لا يحتاج إلى تدبر المترادفات عند البحث في المصطلحات المقيدة التي يستعملها المكشفون نظراً لأن التحكم في الترادف لا بد وأن يكون قد تم في مرحلة التكشيف .

الخلاصة :

من المفيد تلخيص هذا الفصل بمراجعة بعض السياسات والقضايا الجوهرية في التكشيف ، وعقد مقارنة عامة بين ما يسمى عادة بالتكشيف الموضوعي وما يسمى بالفهرسة الموضوعية .

وبناء على الشمول ، فإن الفهرسة الموضوعية بحرصها على استعمال مصطلح واحد أو عدد قليل من المصطلحات التي تلخص مجمل المحتوى المركزي للوثيقة ، هي الأقل شمولاً . وعلى الرغم من أن الشمول عادة ما يحظى بالمناقشة من منظور انتقاء المصطلحات بواسطة البشر ، فإنه من الممكن للتكشيف الجزئي وتكشيف النصوص الكاملة أن يكونا بوجه عام ، أكثر شمولاً ، نظراً لأن أيّاً منهما لا يفاضل بين المصطلحات .

وينطوي التجهيز بواسطة البشر على أعباء مالية ضخمة يتحملها منتج مرصد البيانات ، نظراً لأنه يتطلب المزيد من الاختصاصيين والموارد المادية . أما تكشيف النصوص جزئياً أو كلياً ، فإنه على الرغم من أنه قد لا يكلف منتج مرصد البيانات كثيراً ، فإنه يضع أعباء إضافية في الوقت والجهد على عاتق المستفيد ، وذلك في صياغة استراتيجيات البحث الناجحة .

وبربط ذلك باللغة ، فإن تعيين المصطلحات للتعبير عن المحتوى الموضوعي للوثيقة يستند إلى فكرة التكشيف الموضوعي أو تكشيف المفاهيم اعتماداً على اللغة المقيدة . ويعتمد التكشيف البشري بالاقتراف على مصطلحات المؤلف ، ومن ثم فإنه يمكن أن يسمى بتكشيف الرموز لا تكشيف المفاهيم (Anderson, 1983) إلا أنه عادة ما يكون شبه مقيد (حيث يحظى الاطراد الصرفي بقدر من الانتباه) أما تكشيف النصوص جزئياً أو كلياً ، فهو تكشيف للرموز غير مقيد على الإطلاق .

هذا ، وتتيح العلاقة التكاملية بين حقول الواصفات الانتقائية المقيدة (أساساً) والحقول النصية غير الانتقائية وغير المقيدة ، بعض الفرص . وواقع الأمر أن الحقول النصية غير المقيدة ، والتي غالباً ما تشمل المستخلص ، تؤخذ كما قدمها مؤلفو الوثائق الأولية . ومن الممكن للمستخلص إذا ما عهد به إلى اختصاصي محترف ، أن يصاغ عمداً بما يضمن ثراء التسجيلات الوراقية الناتجة بالترادفات إلا أنه قد تبين لفيدل (Fidel 1986)

أن قلة فقط من منتجي مراصد البيانات يستغلون هذه الفرصة ، بأن يطلبوا ممن يعدون لهم المستخلصات استعمال المصطلحات المترادفة عن قصد .

هذا ، ومن الصعب التعرف على حدود توافر المعلومات الموضوعية الغزيرة ، ومتى يصبح تأثيرها في النظام مجرد مزيد من الشوشرة (مواد غير مرغوبة) لا معلومات إضافية مفيدة (Blair, 1986) . وهذه ولا شك قضية في استرجاع النصوص الكاملة ، حيث يتضح أن اتخاذ جميع الكلمات الواردة في الوثيقة بأكملها وسيلة للتعبير عن المحتوى الموضوعي ، يمكن أن يسفر عن نظم يصعب البحث فيها بشكل فعال . وهناك في الواقع كثير من مقالات النصوص الكاملة المزودة أيضاً بالمصطلحات الكشفية والمستخلصات ، ويركز قطاع رئيس من البحث في استرجاع المعلومات الآن على كيفية تكشف وثائق النصوص الكاملة آلياً . وسوف نتناول هذا الجهد بمزيد من التفصيل في الفصل الحادي عشر .

اللغة في الاسترجاع

لقد ألقت النظرة الشاملة إلى مكونات نظام استرجاع المعلومات في الفصل الثالث الضوء على مركزية اللغة في عملية الاسترجاع ؛ فالمستفيدون يعربون عن حاجتهم إلى المعلومات للوسطاء بمصطلحات لغوية ، ويترجم الوسطاء (من البشر والآلات على السواء) طلب الحصول على المعلومات إلى لغة النظام . ومما لا شك فيه أن اللغة مهمة جداً لأنها تؤثر في اثنين من مكونات نظام الاسترجاع الذي يصوره الشكل رقم (٢) وهما أنشطة المنتجين في إنشاء مراصد البيانات وأنشطة المستفيدين في الإفادة من هذه المراصد.

ولتوفير مقومات الاسترجاع الموضوعي تقوم النظم الحديثة لاسترجاع المعلومات بتجهيز الوثائق والطلبات بطرق تكفل مضاهاة الكلمات والعبارات الواردة في تسجيلات الوثائق بتلك الواردة في الطلبات . ويمكن لتسجيلات الوثائق والطلبات أن تشتمل على المصطلحات المقتطفة من قوائم المصطلحات المحدودة التي قصد بها التعبير عن مجال موضوعي معين (الواصفات من المكنز أو رموز الموضوعات من قائمة رموز الموضوعات) أو تشتمل على بدائل التعبير غير المقيدة المعتمدة على اللغة الطبيعية (المصطلحات من الحصيلة اللغوية (الأجرومية) الخاصة بالمستفيد أو المصطلحات الواردة في الحقول النصية كالعنوان أو المستخلص أو النص الكامل للوثيقة). ويركز هذا الفصل على كل من اللغات المقيدة واللغات غير المقيدة (اللغة الطبيعية) للتعبير عن الموضوعات واسترجاعها ، وخصائص كل من النوعين وما لهما من أثر في عملية البحث .

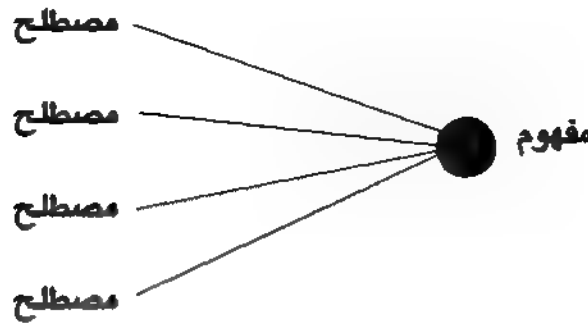
خصائص اللغة واللغة المقيدة :

لإدراك أهداف وخصائص التحكم في اللغة على أفضل وجه ، فإنه ينبغي بادئ ذي بدء إدراك مشكلات الاسترجاع التي تثيرها النظم التي لا تعتمد على اللغات المقيدة .

من الممكن ظهور عدد من المشكلات في نظم الاسترجاع التي يستعمل فيها المكشفون مصطلحات اللغة الطبيعية لوصف المحتوى الموضوعي ، أو في حالة النظم الإلكترونية التي يتم فيها اختزان نصوص الوثائق كاملة أو المستخلصات ، في شكل قابل للتداول بواسطة الآلات (ويستعمل مسئول البحث في مثل هذه النظم اللغة الطبيعية بالطبع) .

الترادف :

من بين المشكلات عدم الاطراد في التعبير عن الموضوع نفسه . وبعبارة أخرى ، يمكن التعبير عن موضوع معين بعدة طرق مختلفة في وثائق مختلفة أو من جانب مكشفين مختلفين . ويعني هذا الموقف الذي يعرف في اللغويات بالترادف ، وجود أكثر من مصطلح واحد للدلالة على موضوع أو مفهوم معين . ويمكن تصوير ذلك على النحو التالي :



فمن الممكن على سبيل المثال التعبير عن الحالة المرضية نفسها من جانب أحد المكشفين بالسل الرئوي PULMONARY TUBERCULOSIS ومن جانب آخر بسل الرئة TUBERCULOSIS OF THE LUNG ، أو التعبير عن العقار metharbital بهذا المصطلح أو بأي من مترادفاته ، مثل : METHYLPHENOBARBITAL أو GEMONIL أو ENDIEMAL (أسماء تجارية) . ومما لا شك فيه إذن ، أنه يتعين على مسئول البحث ، في النظام المعتمد على اللغة غير المقيدة ، أن يأخذ في حسبان جميع الكلمات أو التعبيرات المترادفة ، لكي يسترجع جميع الوثائق المتصلة بموضوع معين . وقد لا يكون العبء على مسئول

البحث ضخماً جداً بالنسبة للكلمات المفردة، إلا أنه يمكن أن يصبح متجاوزاً للحدود في حالة العبارات المترادفة. فمن الممكن على سبيل المثال لفكرة مستوى (إحدى المواد الكيميائية) في الدم أن يتم التعبير عنها بأكثر من طريقة في النص :

| | |
|-----------------------------------|-----------------------|
| blood levels | مستويات الدم |
| serum levels | مستويات مصل الدم |
| blood concentration | تركيز الدم |
| level of ... in the blood | مستوى ... في الدم |
| level of ... in the serum | مستوى ... في مصل الدم |
| concentration of ... in the blood | تركيز ... في الدم |
| levels in the blood | المستويات في الدم |

ولا يمكن بحال لمستول البحث أن يكون قادراً على الإلمام بكل المترادفات .

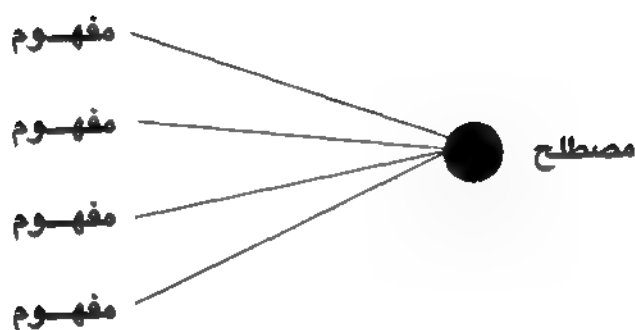
ومن بين الوظائف الأساسية للغة المقيدة التحكم في المترادفات ، أي تحديد أي التعبيرات المترادفة المتعددة يمكن أن يستعمل من جانب المكشفين ومسئولي البحث ، ومن ثم تجنب تشتيت الوثائق المتصلة بالموضوع نفسه تحت مصطلحات مختلفة في النظام . ويمكن لمثل هذا التحكم أن يتم بطريقتين : وتنطوي الطريقة الأولى على اختيار قسم معياري من أقسام الكلام للمصطلحات ، حيث تصاغ عادة في شكل أسماء ، وبذلك نتجنب تشتت الموضوع تحت عدة أوزان صرفية مختلفة (التعليم لا يُعَلِّم أو عُلِّم). وعندما يكون هناك عدة رموز لغوية مختلفة تمام الاختلاف (مثل ONLINE PUBLIC ACCESS CATALOGS/ OPACS) فإن التحكم يتحقق باختيار أحد البدائل المحتملة، أي « المصطلح المفضل » ، ثم الإحالة ، بكلمة « انظر » أو « استعمل » ، إلى المصطلح المفضل من المترادفات التي يمكن للمستفيدين استعمالها في التعامل مع النظام.

ويسفر التحكم في المترادفات ، وهذا هو التعبير الأكثر شيوعاً في الاستعمال للدلالة على هذه العملية ، عن « علاقة تكافؤ » : فالمترادفات الحقيقية مرشحة ولا شك لهذه

العلاقة ، ولكن فيما عدا المختصرات هناك عدد قليل نسبياً من الكلمات الإنجليزية التي تعد مترادفة فعلاً (Ullmann, 1963) . وتشمل المصطلحات التي تعد متكافئة أشباه المترادفات (مباني المكتبات / عمارة المكتبات) والأضداد (كالخشونة في مقابل النعومة) . ومن المؤكد أن هذه التوافق ليست مترادفات بالمعنى اللغوي ، وإنما يتكون كل زوجين منها ، من مصطلحات يمكن اعتبارها متكافئة لأغراض البحث . ويرى كل من ماندرسلوت ودوجلاس وسبايسر (Mandersloot, Douglas and Spicer 1970) على سبيل المثال أن الأضداد التي تعبر عن وجهات نظر مختلفة حول سلسلة الخواص نفسها، يمكن أن تعد قابلة لأن يحل بعضها محل البعض ، أي تستعمل تبادلياً لأغراض الاسترجاع . فمما لا شك فيه أنه من الممكن النظر إلى الخشونة باعتبارها مجرد غياب النعومة والعكس بالعكس ، والمقالة التي تتناول تأثيرات الخشونة على الخواص الديناميكية الهوائية للألواح المعدنية ، تتناول أيضاً تأثيرات النعومة .

الاشتراك اللفظي :

والاشتراك اللفظي مشكلة أخرى ناتجة عن غياب التحكم في اللغة ، ويتمثل في وجود الكلمات المؤتلفة شكلاً والمختلفة معنى ، أي الكلمات المتطابقة في الهجاء والمختلفة في المعنى . والاشتراك اللفظي إحدى الحالات الخاصة للموقف الذي يشار إليه في اللغويات بالغموض ، ومن الممكن تصويره على النحو التالي :



وهكذا يمكن للمصطلح mercury أن يدل على إحدى الشخصيات الأسطورية، وأحد الكواكب السيارة ، وأحد المعادن أو أحد طرز السيارات(*) وعادة ما تميز اللغة

(*) يعني عطارد إلى التجارة والفصاحة والمكر واللصوصية عند الرومان ، كما يعني عطارد أقرب الكواكب السيارة إلى الشمس ، كما يعني معدن الزئبق ، في الوقت نفسه الذي يستعمل فيه كاسم لأحد طرز السيارات . (المترجم) .

المقيدة بين المفردات المؤتلفة شكلاً المختلفة معنى بمحدد للمعنى أو تبصرة بالمجال بين قوسين . فالصيغة عطار (أساطير) تدل على أن هذا المصطلح لا يستعمل إلا للشخصية الأسطورية ، وليس لأي شيء آخر .

و غالباً ما تكون مشكلة الاسترجاع الخاصة بالغموض الناتج عن المفردات المؤتلفة شكلاً المختلفة معنى ، مشكلة نظرية أكثر منها فعلية ؛ فالمفردات التي يمكن أن تكون غامضة في حد ذاتها ، لا تصبح غامضة عندما تستعمل بصحبة مفردات أخرى . فكلما PLANT يمكن أن تكون غامضة ، إلا أنها عندما تستعمل مع STEEL في البحث في أحد نظم الاسترجاع يتوارى الغموض ، ويدرك مسئول البحث أن المصطلح PLANT يتصل بالعمل الصناعي وليس بأي نوع آخر .

البحث الشامل :

والمشكلة الثالثة الناتجة عن غياب التحكم في اللغة هي أنه يتعين على مسئول البحث الإحاطة بجميع المصطلحات المتصلة ببعضها البعض دلاليًا ، واللازمة لإجراء بحث شامل . فلإجراء بحث عن إنتاج الحبوب Cereals في الشرق الأوسط ، فإن مسئول البحث يحتاج إلى الإلمام بجميع المصطلحات التي يمكن أن تدل على الحبوب وتلك التي يمكن أن تدل على الشرق الأوسط . وتجمع اللغة المقيدة مثل هذه المصطلحات المتصلة ببعضها البعض معاً . وإذا كانت اللغة محكمة البنيان فإنها تجمع معاً المصطلحات المتصلة ببعضها البعض هرمياً ، في شكل يدل على علاقة الجنس والنوع ، كما تكشف أيضاً عن العلاقات الدلالية عبر التفريعات الهرمية . ويسمي جاردان (1965) Gardin هذين الضربين من العلاقات بالعلاقة الأفقية Paradigmatic والعلاقة الرأسية Syntagmatic . وتتسم العلاقة الأفقية بالثبات ، فهي علاقة قائمة دائماً ، ومن أمثلتها المصطلحات «اليومنيوم» و «ماغنسيوم» و «المعادن الخفيفة» . أما العلاقة الرأسية فمؤقتة ، فهي علاقة لا توجد إلا في مواقف معينة فقط . « فالاليومنيوم » يمكن أن يكون على علاقة «بقدر الطهي» ، إلا أنه لا يتصل بقدر الطهي على الدوام ، كما أن قدر الطهي لا تتصل دائماً بالاليومنيوم .

وظائف اللغة المقيدة :

والوظائف الرئيسة للغة المقيدة هي :

١ - ضمان التعبير عن الموضوع بشكل مطرد ، ومن ثم تجنب التششتت الموضوعي في

مرحلتي المدخلات (التكتيف) والمخرجات (البحث) ، وذلك بالتحكم في المترادفات وأشباه المترادفات ، فضلاً عن التمييز بين المصطلحات المؤتلفة شكلاً المختلفة معنى .

٢ - تيسير إجراء عمليات البحث العريضة (الشاملة) بتجميع المصطلحات المتصلة ببعضها البعض دلاليًا ، سواء كانت العلاقات بينها أفقية أو رأسية ، وذلك بطريقة معينة .

فأهداف اللغة المقيدة إذن هي تحقيق علاقة تطابق أو تكافؤ تام بين المصطلحات اللغوية والأشياء أو الأفكار أو المفاهيم التي تدل عليها ، وتحديد ما بين المصطلحات أو المفاهيم التي تدل عليها وبعضها البعض من علاقات أفقية أو رأسية .

اختيار المصطلحات :

الخطوة الأولى في بناء لغة مقيدة هي اختيار ما يمكن أن تشتمل عليه من مصطلحات. ويحدد لانكستر (1986) Lancaster معالم أهم سبل تجميع المصطلحات، بما في ذلك المسوغ الأدبي (حينما تلتقط المصطلحات من نوعية الإنتاج الفكري المزمع تكتيفه) ومسوغ الإفادة (حيثما تأتي المصطلحات من المستفيدين المحتملين من مرفق المعلومات) . وهناك قضيتان لهما تأثيرهما في التكتيف والبحث ، تثيران المشكلات بوجه خاص في اختيار المصطلحات :

١ - المستوى المناسب للتخصيص في اللغة .

٢ - مدى ارتباط الكلمات ببعضها البعض لتشكيل مصطلحات كشفية أكثر تعقيداً وأكثر تخصيصاً .

ولمسوغ الإفادة أهميته الخاصة في تحديد مستوى تخصيص المصطلحات التي يقع عليها الاختيار . فمن الممكن إعداد قائمة تشتمل على عدة مئات من فصائل الكلاب ، والتي يمكن لأسمائها أن يكون لها مسوغها الوراقي، ولكن إذا كان المستفيدون من نظام معين لا يحتاجون على الإطلاق إلى ما هو أكثر تخصيصاً من « الكلاب » فإنه ليس هناك من مبرر مقبول لتطوير هذا القطاع من اللغة بمثل هذا القدر الكبير من التحليل أو التفصيل . ويعني ذلك أنه يتعين توافر قدر كبير من المعلومات حول المستفيدين

المحتملين من النظام وحول نوعيات ما يمكن أن يتقدموا به من طلبات ، وذلك في متناول المسئول عن إعداد اللغة المقيدة . وهذه نقطة ركز عليها وبقوة كل من سورجل Soergel (1974) ولانكستر (1972, 1986) .

وبمجرد أن يقرر المسئول عن إعداد اللغة المقيدة إدخال مفهوم أو موضوع ما لأن له ما يسوغه وراقياً أو ما يسوغه من وجهة نظر المستفيدين ، فإنه ينبغي أن يقرر ما إذا كان من الممكن التعبير عن هذا الموضوع على أحسن وجه ، بعبارة مركبة أم بتجزئ العبارة إلى عنصرين أو أكثر . لننظر في وثيقة افتراضية حول اضطرابات التعلم لدى أطفال ما قبل المدرسة ، ويمكن لهذه الوثيقة أن تكشف في مستوى من الشمول منخفض جداً (انظر الفصل الخامس) .

وبناء على نوعية اللغة المستعملة ، يمكن تكشيف هذه الوثيقة باستعمال مصطلح واحد أو مصطلحين أو أربعة مصطلحات ، كما هو موضح في الشكل رقم (٢٧) . وبناء على كيفية تجميع المفاهيم في مرحلة تصميم اللغة ، فإنه من الممكن النظر إليها باعتبارها تنتمي إلى فئة واحدة أو عدة فئات محتملة من الوثائق ، كما تعبر عن ذلك أشكال فن المصاحبة^(١) . ويوضح هذا المثال الفرق بين الربط المسبق والربط اللاحق في استرجاع المعلومات . فكل من الوثائق والطلبات تتكون من مفاهيم فرعية أو أوجه ، « كاللغة » ، و«الاضطرابات» ، و« ما قبل المدرسة » ، و« الأطفال » . وفي ظل الربط المسبق يتم الجمع بين هذه المفاهيم بطريقة ما والتعبير عنها بعبارة واحدة مركبة أو أكثر من عبارة (كرأس رئيسي مصحوب برأس فرعي مثلاً) . ولا يمكن الربط بين المفاهيم أو الأوجه التي ارتبطت ببعضها مسبقاً على هذا النحو ، صراحة في مرحلة البحث . أما في ظل الربط اللاحق ، فإنه يستعمل في التعبير عن المحتوى الموضوعي للوثيقة مصطلحات تدل على مفاهيم بسيطة ، ينبغي بعد ذلك الربط بينها صراحة بالروابط البولينية ، في مرحلة البحث للتعبير عن الارتباطات الدلالية المركبة . وعلى ذلك ، فإن الربط المسبق والربط اللاحق أنشطة تنطبق على اللغة المستعملة في تكشيف الوثائق وعلى العمليات التي ينبغي أن تتم للبحث في مرصد البيانات .

(١) سوف نتناول أشكال فن Venn diagrams بمزيد من التفصيل في الفصل الثامن .

التعبير المنطقي عن الفئات التي
ينبغي البحث عنها لاسترجاع
المعلومات عن الموضوع

المصطلح / المصطلحات المستعملة

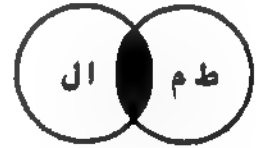
١. ربط مسبق تام (لا مجال
لربط بين المصطلحات
أثناء إجراء البحث).

اضطرابات اللغة في
أطفال ما قبل المدرسة



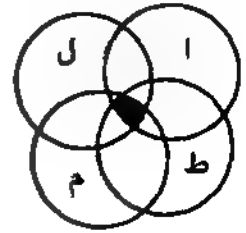
ب. ربط مسبق جزئي وربط
لاحق جزئي (ينبغي
الربط بين مصطلحين
لاسترجاع المعلومات عن
الموضوع).

اضطرابات اللغة
أطفال ما قبل المدرسة



جـ ربط لاحق تام (تم
تحليل جميع الأوجه
إلى كلمات مفردة وكلها
يمكن الجمع بينها
لاسترجاع المعلومات
عن الموضوع).

اللغة
الاضطرابات
ما قبل المدرسة
الأطفال



الشكل رقم (٢٧) : مقارنة بين نظام الربط المسبق ونظام الربط اللاحق في التكشيف.

ويميز لانكستر (1972) Lancaster بين اللغات الحصرية واللغات التحليلية التركيبية. وتقدم اللغة الحصرية المطلقة، ببساطة قائمة بالمصطلحات، لا يمكن الربط بينها للدلالة على موضوع أكثر تعقداً. أما اللغات التحليلية التركيبية، فعلى العكس، يمكن أن تسمح بالربط بين المصطلحات لتشكيل مصطلحات أخرى أكثر تعقداً، سواء في مرحلة التكشيف أو في مرحلة البحث. والفارق بين الحصري والتحليلي التركيبي قد لا يكون واضحاً تماماً؛ فقائمة رءوس موضوعات مكتبة الكونجرس Library of Congress Subject Headings التي تحصر عدداً كبيراً من المصطلحات المركبة مسبقاً، تقدم أيضاً بعض التيسيرات اللازمة لبناء مصطلحات جديدة بإضافة رءوس فرعية إلى الرءوس

الرئيسية . كما أن المكانز على الرغم من أنها مصممة في الأساس للاستخدام في تطبيقات الربط اللاحق ، تشتمل على كثير من العبارات التي تدل على الربط المسبق بين الموضوعات .

ورقم (١) في الشكل رقم (٢٧) نموذج للغة الحصرية المطلقة الخاصة بالربط المسبق ، بينما رقم (ج) في الشكل نفسه تحليلي تركيبى ولاحق الربط . أما رقم (ب) في الشكل رقم (٢٧) فيدل على درجة ما من الحصر والربط المسبق باستعماله للعبارات ، كما يسمح أيضاً ببعض التحليل والتركيب أو الربط اللاحق ، بجعل هذه العبارات قابلة للربط في مرحلة البحث .

ويلخص الشكل رقم (٢٨) أوجه الاختلاف الأساسية بين الربط المسبق والربط اللاحق . وربما يبدو أنه كلما مالت اللغة نحو الربط اللاحق فإن الأمر يتطلب المزيد من المصطلحات الدقيقة لتغطية الموضوعات في التكشيف والبحث على السواء . وهذا أمر مضلل في الواقع ويوضح المخاطر المرتبطة بالتعميم اعتماداً على مثال واحد . والواقع أنه من الممكن للكلمات المفردة المستعملة كمداخل كشفية في رقم (ج) في الشكل رقم (٢٧) أن تكون عناصر في تكوين العديد من العبارات المختلفة ؛ حيث يمكن بناء العديد من العبارات اعتماداً على عدد قليل من الكلمات ، بالطريقة نفسها التي يمكن بها تكوين عدد كبير من الكلمات اعتماداً على عدد قليل من الحروف الهجائية . ومن ثم ، فإنه من الممكن ، بالنسبة لمجال موضوعي معين ، للغة المفردة في الربط المسبق أن تشتمل على عدد من المصطلحات أكثر بكثير مما تشتمل عليه لغة الربط اللاحق ، ولهذا ، فإن لغات الربط المسبق تشتمل على تكرار للمفاهيم أكثر مما تشتمل عليه لغات الربط اللاحق (Millstead, 1984) . ومن الممكن توضيح ذلك بالمثال التالي :

| <u>المنتج</u> | <u>عملية التجهيز</u> |
|---------------|----------------------|
| التفاح | التجفيف |
| الكمثرى | التعليب |
| ... إلخ | ... إلخ |

ويمكن للغة الربط المسبق أن تشتمل على كل مصطلح من المصطلحات المفردة

(التفاح ، والكمثرى ، والتجفيف ، والتعليب) فضلاً عن جميع المصطلحات التي تربط بين كل من الفاكهة وعملية التجهيز (تجفيف التفاح ، وتعليب التفاح ، وتجفيف الكمثرى ، وتعليب الكمثرى) . وينطوي ذلك على حشو ويجعل لغة الربط المسبق أطول بكثير من لغة الربط اللاحق التي تناظرها ، والتي تشتمل فقط على المصطلحات المفردة ، بينما يتم التعبير عن العلاقات الأكثر تعقداً في مرحلة البحث .

| عدد المصطلحات في الكشف / | الربط المسبق | الربط اللاحق |
|----------------------------|--------------|---|
| الطلب | أقل | أكثر |
| عدد المصطلحات في اللغة | أكثر | أقل |
| تكرار المفاهيم | أكثر | أقل |
| مدى تخصيص المصطلحات | أكثر | أقل |
| المرونة في ارتباط المفاهيم | أقل | أكثر |
| نوعية الملف | ورقي | إلكتروني (كانت الأشكال المبكرة تشمل الورق أو المصغرات الفلمية) |
| الأبعاد | أحادي البعد | متعدد الأبعاد |
| العلاقات الزائفة | أقل | أكثر |

الشكل رقم (٢٨) : أوجه الاختلاف بين كشف الربط المسبق وكشف الربط اللاحق .

وتبدو لغات الربط المسبق أكثر تخصيصاً لأنها تشتمل على مصطلحات أطول وأكثر تفصيلاً من لغات الربط اللاحق . وعلى الرغم من أن كلا من لغات الربط المسبق ولغات الربط اللاحق تسمح بإجراء عمليات البحث عن الموضوعات المخصصة ، فإن هذه الموضوعات يتم التعبير عنها تحديداً من البداية بالنسبة لنظم الربط المسبق ، وتشق أثناء عملية البحث بالنسبة لنظم الربط اللاحق . وتميل نظم الربط اللاحق ، على المدى الطويل لكفالة درجة من التخصيص أعلى مما تكفله نظم الربط المسبق ، نظراً لأنها تسمح بالتعبير بمرونة ويسر عن العلاقات الجديدة بين المفاهيم أو الموضوعات .

وتكاد نظم الاسترجاع المعتمدة على الملفات المطبوعة ، كالفهارس البطاقية والكشافات المطبوعة ، تكاد تكون بلا استثناء من نظم الربط المسبق . وغالباً ما يكون

المستفيدون من هذه النظم بحاجة لأن يفكروا في مدخل واحد فقط ، إلا أن هذا المدخل ينبغي أن يكون المدخل الذي حدده المكشف ، كما أنه ينبغي البحث في الملف ، دون خطأ ، عن طريق الكلمة الأولى في الخيط . ومن الممكن التخفيف من حدة هذا الجمود إلى حد ما بإدراج الإشارات الوراقية تحت مختلف تباديل العناصر نفسها التي تتكون منها المصطلحات ، كما في نظام بريسييز PRECIS (Austin, 1984) إلا أن ذلك يضيف بشكل ملحوظ إلى حجم الملف ، ومن ثم إلى تكلفته . أما الربط اللاحق ، فعلى العكس ، يسفر عن عدد كبير من التوافق بين المصطلحات ، إلا أنه لا يمكن إلا للحاسب الآلي أن يكفل القوة الكافية والدقة ، في تحقيق المرونة في الربط والتقاطع بين المصطلحات عند إجراء عمليات البحث المركبة . ويسمى لانكستر (Lancaster (1986) الضربين من الملفات اللذين أشرنا إليهما توأ بالملف أحادي البعد (الورقي) والملف متعدد الأبعاد (الإلكتروني) بينما يسميهما برنيير (Bernier (1956) بالملف الجامد nonmanipulative (الورقي) والملف المرن manipulative (الإلكتروني) . أما سورجل (Soergel (1974) فيفضل المصطلحين التركيب المسبق "pcombination" والتركيب اللاحق

"postcombination" . ورغم كثرة ما تتمتع به نظم الربط اللاحق من مزايا ، فإن هناك بعض المشكلات التي تكتنفها : فمن الممكن لمظاهر الغموض النظمي والدلالي ، وخاصة في نظم الكلمات الأحادية ، أن تكون سبباً في استرجاع وثائق لا تتصل بموضوع الطلب على الإطلاق . فانظر في المجموعة التالية من المصطلحات التي استعملت جميعها في كشف أحد التقارير :

| | |
|------------|-----------|
| ALUMINUM | أليومنيوم |
| COPPER | نحاس |
| WELDING | لحام |
| CLEANING | تنظيف |
| ULTRASONIC | فوقسمعي |

ويتناول التقرير تصنيع المكونات الإلكترونية ، وتنطوي إحدى العمليات على لحام

الأليومنيوم وأخرى على تنظيف النحاس بالموجات فوقسمعية . ويمكن لهذا التقرير أن يُسترجع للأسف استجابة لطلبات المعلومات حول لحام النحاس وحول تنظيف الأليومنيوم وحول اللحام بالموجات فوقسمعية ، على الرغم من أنه لا يتصل بموضوع أي منها. فقد حدثت مصاحبات مزيفة أو خاطئة أو ارتباطات خاطئة بين المصطلحات ، فاسترجعت الوثائق التي كشفت تحت مصطلحين أو أكثر من المصطلحات التي حددها مسئول البحث ، أو كانت تشتمل على هذه المصطلحات ، على الرغم من أنها لا تتصل ببعضها البعض في الأساس في الوثائق أو في تسجيلاتها .

وهناك مظهر آخر محتمل للغموض يمكن أن يسمى بالعلاقة غير الصحيحة بين المصطلحات. انظر إلى طلب للمعلومات أو للإنتاج الفكري حول قلق الانفصال ، أي قلق الطفل الذي انفصل عن أمه ، حيث يمكن للبحث عن هذا الموضوع أن يسفر عن استرجاع تقرير تم تكشيفه على النحو التالي :

| | |
|----------|----------|
| MOTHER | الأم |
| CHILD | الطفل |
| ANXIETY | القلق |
| ILLNESS | المرض |
| HOSPITAL | المستشفى |

إلا أن هذا التقرير لا يتناول قلق الانفصال، وإنما يتناول قلق الأم على الطفل المريض الذي ينبغي أن يدخل المستشفى . وليس هذا بالربط الخاطئ لأن المصطلحات «القلق» و « الأم » و « الطفل » تتصل جميعها ببعضها البعض مباشرة في الوثائق المسترجعة. ومن ثم فإن هذا الموقف يعد مثالاً جيداً للعلاقة غير الصحيحة بين المصطلحات، وهو موقف يحدث عندما تكون المصطلحات التي أدت إلى استرجاع الوثيقة مرتبطة ببعضها البعض على نحو يختلف عما يريده المستفيد ؛ فالأم في هذا التقرير هي القلقة وليس الطفل .

وهذه أمثلة بسيطة جداً لأنواع العلاقات الغامضة الزائفة التي يمكن أن ترد في نظام ما لاسترجاع المعلومات . ويؤدي هذا النوع من مشكلات الاتصال إلى استرجاع وثائق لا تتصل بموضوع طلب المعلومات ، أي حدوث « شوشرة » في النظام ، بالمعنى الاتصالي للكلمة . ومما لا شك فيه أنه كلما زاد عدد المصطلحات المستعملة في تكثيف الوثيقة ، ازدادت احتمالات الارتباطات الزائفة والعلاقات غير الصحيحة بين المصطلحات . ولما كانت أعداد كبيرة من النظم الإلكترونية يتم التكثيف فيها عادة عند مستوى ما بين عشرة مصطلحات وثلاثين مصطلحاً للوثيقة ، فإن هذه المشكلات يمكن أن تكون حادة فعلاً .

وهناك بعض سبل تجنب الارتباطات الزائفة والعلاقات غير الصحيحة بين المصطلحات الكشفية (كما هو موضح في الشكل رقم (٢٧)) كقلق الانفصال وقلق الأمومة مثلاً، أو ، وهو ما يتصل بذلك أيضاً ، إتاحة إمكانية استعمال شكل ما من الرؤوس الفرعية، أي استعمال أحد المصطلحات كتقسيم فرعي لمصطلح آخر، كما في المثال : «أليومنيوم/لحام» . وهناك طريقة أخرى للتغلب على أوجه الغموض هذه ، كانت من الخصائص المميزة للنظم خلال الستينيات ومطلع السبعينيات من القرن العشرين، وهي استعمال الروابط والأدوار . فمن الممكن للروابط أن تحد من عدد الارتباطات الزائفة في نظام الاسترجاع بربط المصطلحات المتصلة ببعضها البعض في الوثيقة معاً، وعزل (تركها بلا ربط) المصطلحات غير المتصلة ، كما في هذا المثال (Lancaster, 1972):

| | |
|-----------|-------|
| أليومنيوم | (أ) |
| تنظيف | (ب) |
| نحاس | (ب) |
| فوقسمعي | (ب) |
| لحام | (أ) |

ونظراً للروابط المستعملة كان من الممكن استرجاع هذه الوثيقة استجابة للطلبات الخاصة بلحام الأليومنيوم أو تنظيف النحاس بالموجات فوقسمعية ، وليس استجابة للطلبات الخاصة بلحام النحاس أو تنظيف الأليومنيوم بالموجات فوقسمعية .

أما الأدوار ، وهي وسائل أكثر تعقداً ، فتفسر العلاقة الدلالية الفعلية بين المصطلحات الكشفية (Lancaster 1972) :

| | |
|-----------------|-----|
| تصميم | |
| الطائفة | (٤) |
| الحاسبات الآلية | (٢) |

حيث يدل (٤) على المستهدف بالفعل ، أو المريض أو المتلقي ، ويدل (٢) على الأداة أو العامل أو وسيلة التنفيذ . ومن الممكن استرجاع هذه الوثيقة استجابة لطلب حول تصميم الطائرات بالحاسبات الآلية ، لا استجابة لطلب حول تصميم الحاسبات الآلية . وكثير من العلاقات الغامضة نظرياً ليست غامضة في التطبيق العملي ؛ فمن الممكن للمصطلحات « انجلترا » و « الضأن » و « نيوزيلاندة » و « تصدير » أن تدل على تصدير الضأن من انجلترا إلى نيوزيلاندة ، إلا أن الموقف العكسي هو الأقرب إلى الحقيقة . وسوف نتناول مشكلة الغموض وسبل تجنبها بمزيد من التفصيل في الفصل العاشر .

وما زال قرار مدى استعمال الربط المسبق في اللغة المقيدة بلا حسم حتى الآن في الإنتاج الفكري . وقد اقترح جونز (Jones (1971) حلاً لمجال موضوعي بعينه (تقنيات المطاط) ، حيث يقدم بعض القواعد النظامية والدلالية التي يمكن بناء عليها أن يتقرر متى يتم الفصل بين عناصر المصطلحات المركبة ومتى يتم الإبقاء عليها كما هي . كذلك تحظى القضية بالتغطية في المواصفات القياسية والقواعد الإرشادية الخاصة ببناء المكانز ، ونخص منها بالذكر *Guidelines for the Establishment and Development of Monolingual Thesauri* (British Standards Institution, 1979) والتي انتقدها جونز (1981) Jones فيما بعد . و *Unesco Guidelines for the Establishment and Development of Monolingual Thesauri* (Unesco, 1981) والتي تتطابق إلى حد بعيد مع المواصفة القياسية (ISO 2788 (International Organization for Standardization, 1986) والمواصفة القياسية البريطانية (B. S. 5723 (British Standard Institution, 1987)

عادة ما تستعمل النظم الحديثة لاسترجاع المعلومات لغة مقيّدة لتكشيف مقالات الدوريات والتقارير التقنية وغيرها من الوثائق التي نجدها في مراصد البيانات الوراقية على الخط المباشر، وما يناظر هذه المراصد من كشافات مطبوعة، فضلاً عن بعض مراصد بيانات النصوص الكاملة. ويمكن للغة المقيّدة أن تتخذ عدة أشكال، بما في ذلك المكانز وقوائم رموس الموضوعات، وخطط التصنيف. والمكانز هي الأكثر انتشاراً في الاستخدام، إلا أن مكانز استرجاع المعلومات تختلف كثيراً عن المكانز التقليدية من نوعية مكانز روجيه Roget.

المكانز:

المكانز عادة عبارة عن قائمة هجائية بالمصطلحات المتخصصة في مجال موضوعي معين، والتي يمكن أن تستعمل في التكشيف والبحث عن الإنتاج الفكري. وهو يكفل التحكم في المترادفات، ويميز بين المصطلحات المؤتلفة مبنى والمختلفة معنى. كما يجمع المصطلحات المتصلة ببعضها البعض معاً. وأمامنا المثال البسيط لمداخل المكانز المنتقاة في الشكل رقم (٢٩)، حيث تم « التحكم في الكلمات التي تعد متقاربة بما فيه الكفاية في المعنى، بحيث تعد مترادفات، فوق الاختيار على إحداها تتم الإشارة إليها تحت الكلمات الأخرى بالإحالة أو التوجيه « استعمل ». وقد لا يستعمل المكشف المصطلح «حبوب» إلا أنه ينبغي أن يستعمل بدلاً منه المصطلح « غلال » وبذلك يتجنب تشتت الموضوعات المرتبطة ببعضها البعض. أما الكلمات المؤتلفة مبنى والمختلفة معنى (المشترك اللفظي) فيتم التمييز بينها وفصلها عن بعضها البعض بتبصرة المجال: (PLANTS (BOTANY و (PLANTS (INDUSTRY. ويتم الربط بين المصطلحات المتصلة ببعضها البعض دلاليّاً بطريقتين: فالكلمات المتصلة ببعضها البعض اصطلاحياً كالعلاقة بين الجنس والنوع، يتم الإشارة إليها « بالمصطلحات الأعرض » و« المصطلحات الأضيق ». فالمصطلح « غلال GRAIN » ورد تحته مصطلح أعرض منه، أي الجنس الذي ينتمي إليه وهو « المحاصيل CROPS » ومصطلحات أضيق منه، أي الأنواع المتفرعة عنه، وهي على وجه التحديد أنواع الغلال كل على حدة. أما المصطلحات المتصلة دلاليّاً «بالغلال» وعلى نحو يختلف عن علاقة الجنس بالأنواع، كالعمليات الزراعية أو الصناعية مثلاً المتصلة « بالغلال » فقد وردت كمصطلحات

متصلة . كما ينبغي أن نلاحظ أيضاً أن الإحالات تتم بشكل تبادلي : « فالأذرة » يظهر « الغلال » كمصطلح أعرض منه ، ومن ثم ينبغي أن يورد « الغلال » « الأذرة » كمصطلح أضيق منه . ولما كان المصطلح CEREALS يحال منه إلى GRAIN (أي CEREALS GRAIN) فإن GRAIN ينبغي أن يظهر الإحالة من CEREALS (أي أنه يستعمل بدلاً منه CEREALS) .

| | | |
|--------------------------------|---|----------------------------------|
| Barley | الشعير | |
| Broader term : Grain | مصطلح أعرض : الغلال | |
| Cereals | الحبوب | |
| Use : Grain | استعمل : الغلال | |
| Corn | الأذرة | |
| Broader Term : Grain | مصطلح أعرض : الغلال | |
| Foctories | المصانع | |
| Used for : Plants (industry) | يستعمل بدلاً منه : المواقع (الصناعية) | |
| Grain | الغلال | |
| Used for : Cereals | يستعمل بدلاً منه : الحبوب | |
| Broader term : Crops | مصطلح أعرض : المحاصيل | |
| Narrower Terms : Barley | الشعير (*) مصطلحات أضيق الدقيق | |
| Related terms : Flour | مصطلحات متصلة | |
| الأذرة (*) | Corn | |
| الأذرة (*) | Maize | النباتات (علم النبات) (*) |
| الشوفان | Oats | Plants (botany) |
| الشعير (*) | Rye | المواقع (الصناعية) (*) |
| القمح | Wheat | Plants (industry) |
| الطواحين | Flour mills | استعمل : المصانع Use : Factories |
| الحصاد | Harvesting | |
| الطحن | Milling | |
| الندراس | Threshing | |

الشكل رقم (٢٩) : بنية المكنز .

(*) يؤكد ذلك غياب التطابق التام بين اللغة العربية واللغة الإنجليزية ، وهذا ما ينبغي مراعاته عند التفكير في ترجمة المكانز الإنجليزية إلى العربية . (المترجم) .

وعلى ذلك فإن مسئول البحث ، وكذلك المكشف ، يجد أمامه صورة مكتملة لجميع المصطلحات الواردة في اللغة ، والتي اعتبرت متصلة بمصطلح Grain . وبإمكان المكنز الحيلولة دون تشتت الوثائق المتصلة ببعضها البعض تحت المصطلحات المترادفة ، وكذلك التمييز بين المصطلحات المؤتلفة مبنى المختلفة معنى ، فضلاً عن مساعدة مسئول البحث على تحقيق الشمول في البحث عن مجال موضوعي معين . كذلك يمكن لمسئول البحث الذي يرغب في توسعة الطلب بمصطلحات تدل على أنواع المحاصيل ، أن يفعل ذلك بمجرد إدخال كل المصطلحات تلك الواردة كمصطلحات أضيق ، في استراتيجية البحث . كذلك قد يتراءى لمسئول البحث توسعة البحث في اتجاهات دلالية أخرى ، بإدخال المصطلحات المتصلة ، وبإمكان المكنز الذي أحسن إعداداته أن يميز هذه المصطلحات صراحة من المصطلحات الخاصة بعلاقة الجنس بالنوع . ويساعد هذا كلاً من المكشف ومسئول البحث في اختيار المصطلحات الأكثر من غيرها ملاءمة لموقف معين.

رءوس الموضوعات :

وقائمة رءوس الموضوعات شكل آخر من أشكال اللغات المقيدة ، وهي نوعية اللغة التي تستعمل تقليدياً لتوفير المدخل الموضوعي للكتب وغيرها من الإنتاج الفكري في فهارس المكتبات . ويوضح الشكلان رقماً (٣٠) و (٣١) مقتطفين من رءوس الموضوعات . وتلتزم قوائم رءوس الموضوعات بالمبادئ الأساسية للغة المقيدة ، من التحكم في المترادفات ، والتمييز بين المصطلحات المؤتلفة مبنى والمختلفة معنى ، والبنية الرابطة المعبرة عن العلاقات الدلالية بين المصطلحات . ويتم التحكم في المترادفات والتمييز بين المصطلحات المؤتلفة مبنى المختلفة معنى في قوائم رءوس الموضوعات ، كما يتم في المكانز تماماً ؛ حيث يتم تحديد المصطلح المفضل من بين مجموعة من المصطلحات المتناظرة (التحكم في المترادفات) كما تستعمل المحددات الواردة بين الأقواس (التمييز بين المصطلحات المؤتلفة مبنى والمختلفة معنى) . كذلك يمكن لقائمة رءوس الموضوعات أن تميز بين علاقات الجنس والنوع ، وغيرها من العلاقات الدلالية في بنيتها الترابطية ، كما في الشكل رقم (٣٠) المقتطف من قائمة رءوس الموضوعات الطبية *Medical Subject Headings (Me SH)* كما يمكن لهذا التمييز أن تضيع معاملة الواضحة كما في الشكل رقم (٣١) المقتطف من قائمة رءوس موضوعات مكتبة الكونجرس *Library of Congress Subject Headings (LCSH)*

أ. البنية الشجرية

MEDICINE

| | |
|------------------------|--------------------|
| ADOLESCENT MEDICINE | G2.403.28 |
| BEHAVIORAL MEDICINE | G2.403.90 |
| COMMUNITY MEDICINE | G2.403.220 |
| EPIDEMIOLOGY | G2.403.290 |
| GENETICS, MEDICAL | G2.403.388 |
| GERIATRICS | G2.403.398 |
| MILITARY MEDICINE | G2.403.458 |
| NAVAL MEDICINE | G2.403.478 |
| SUBMARINE MEDICINE | G2.403.478.508 |
| PSYCHIATRY | G2.403.642 |
| ADOLESCENT PSYCHIATRY | G2.403.642.80 |
| BIOLOGICAL PSYCHIATRY | G2.403.642.100 |
| CHILD PSYCHIATRY | G2.403.642.130 |
| COMMUNITY PSYCHIATRY | G2.403.642.150 |
| PREVENTIVE PSYCHIATRY | G2.403.642.150.580 |
| FORENSIC PSYCHIATRY | G2.403.642.208 |
| GERIATRIC PSYCHIATRY | G2.403.642.260 |
| MILITARY PSYCHIATRY | G2.403.642.508 |
| SOCIAL MEDICINE | G2.403.720 |
| SPECIALTIES, MEDICAL | G2.403.776 |
| ALLERGY AND IMMUNOLOGY | G2.403.776.30 |
| DERMATOLOGY | G2.403.776.158 |
| FAMILY PRACTICE | G2.403.776.230 |
| INTERNAL MEDICINE | G2.403.776.409 |
| CARDIOLOGY | G2.403.776.409.163 |
| ENDOCRINOLOGY | G2.403.776.409.323 |
| GASTROENTEROLOGY | G2.403.776.409.440 |
| HEMATOLOGY | G2.403.776.409.543 |
| MEDICAL ONCOLOGY | G2.403.776.409.708 |
| NEPHROLOGY | G2.403.776.409.752 |
| RHEUMATOLOGY | G2.403.776.409.808 |
| PEDIATRICS | G2.403.776.671 |

ب. العرض الهجائي

CHILD, PRESCHOOL

M1.471.392.448

age 2-5 yr; IM only as psychol & sociol or social entity: Manual 18.5.11,34.10; NIM as check tag: Manual 18.5+; no qualif for IM but psychol of the preschool child = CHILD PSYCHOLOGY (IM) & check the tag CHILD, PRESCHOOL: Manual 34.12

66

CHILD PSYCHIATRY

F4.96.544.193

G2.403.642.130

G2.403.790.600.258

SPEC: SPEC qualif; not for mental disord in children (= MENTAL DISORDERS (IM) + CHILD or other child check tag (NIM): Manual 27.9

CHILD PSYCHOLOGY

F4.96.628.193

SPEC but also the way a normal child thinks & acts; GEN only: consider also / psychol with specific child terms (e.g., CHILD, ABANDONED/psychol, ONLY CHILD/ psychol): SPEC qualif; Manual 27.12,34.12; check also tag CHILD or specific

X INFANT PSYCHOLOGY

XR CHILD DEVELOPMENT

الشكل رقم (٣٠) : نموذج طريقتي العرض في رموس الموضوعات الطبية .

Child psychiatry (Indirect) (RJ499)

Here are entered works on the clinical and therapeutic aspects of mental disorders in children. Descriptive works on mental disorders of children are entered under Child psychopathology. Works on mentally ill children themselves are entered under Mentally ill children.

sa Adolescent psychiatry

Autism

Child development deviations

Child guidance clinics

Child mental health

Child psychology

Child psychopathology

Child psychotherapy

Children of the mentally ill

Cognition disorders in children

Hysteria in children

Infant psychiatry

Mentally handicapped children

Mentally ill children

Psychomotor disorders in children

Psychoses in children

Schizophrenia in children

School phobia

Sleep disorders in children

x Children - Mental disorders

Pediatric psychiatry

Psychiatry, Child

xx Child mental health

Child mental health services

Child psychology

Child psychopathology

Pediatric neurology

Psychiatry

Child psychiatry (May Subd Geog) [RJ499]

Here are entered works on the clinical and therapeutic aspects of mental disorders in children. Descriptive works on mental disorders of children are entered under Child psychopathology. Works on mentally ill children themselves are entered under Mentally ill children.

UF Children-Mental disorders

Pediatric psychiatry

Psychiatry, Child

BT Child mental health services

Pediatric neurology

Psychiatry

RT Child mental health

Child psychology

Child psychopathology

NT Adolescent psychiatry

Child development deviations

Child Psychotherapy

Children of the mentally ill

Cognition disorders in children

Hysteria in children

Infant psychiatry

Interviewing in child psychiatry

Mentally handicapped children

Mentally ill children

Psychomotor disorders in children

Psychoses in children

Schizophrenia in children

Sleep disorders in children

الشكل رقم (٣١) : مقارنة بين عرض المصطلحات في الطبعة العاشرة والطبعة الحادية عشرة من قائمة
رعوس موضوعات مكتبة الكونجرس .

وفي قائمة رءوس الموضوعات الطبية يعبر الترتيب الهجائي للمصطلحات صراحة عن العلاقات التي لا تدخل ضمن علاقات الجنس بالنوع (فعلم نفس الطفل CHILD PSYCHOLOGY قد ورد بشكل يدل على ارتباطه XR بنمو الطفل CHILD DEVELOPMENT) كذلك تسجل القائمة أمام كل مصطلح رقم أو أرقام التفريعات الهرمية التي يظهر فيها ، وبذلك تقود أو ترشد إلى البنى الشجرية . وفي هذه اللغة إذن يتولى العرض الهرمي (الشجري) مهمة التعبير عن علاقات الجنس بالنوع ، بينما يتعامل العرض الهجائي مع مشكلة الترادف (فعلم نفس الطفل CHILD PSYCHOLOGY يحال إليه من - برمز x - INFANT PSYCHOLOGY) وكذلك مع العلاقات خلاف علاقات الجنس بالنوع .

وكانت الترميزات المستعملة للدلالة على البنية الترابطية ، في الطبقات القديمة من قائمة رءوس موضوعات مكتبة الكونجرس ، تختلف عن تلك التي نجدها في المكانز (انظر القائمة الواردة في الجانب الأيسر للشكل رقم (٣١)) . فقد كانت إحالة « انظر » بدلاً من « استعمل » تحيل المستفيد من أحد العناصر الواردة كمدخل إلى المصطلح المفضل (فكان من الممكن على سبيل المثال ، لمدخل لم يظهر في الشكل رقم (٣١) أن يحيل من «الأطفال - اضطرابات عقلية Children - Mental disorders» إلى « الطب النفسي للطفل Child psy- chiatry ») ، كما كانت x (« انظر من ») بدلاً من « يستعمل بدلاً من Used for » هي البديل المستعمل تحت المصطلح المفضل للدلالة على مصطلحات المداخل المناظرة . كذلك كانت الرموز sa (انظر أيضاً see also) و xx (« انظر أيضاً من see also from ») تستعمل معاً للدلالة على كل من علاقات الجنس بالنوع (الأعرض / الأضيق) والعلاقات غير التفريعية (الترابطية) . ويرى دايكسترا (Dykstra 1988 b) أن الطبعة الحادية عشرة من قائمة رءوس موضوعات مكتبة الكونجرس قد أعدت لتكون مشابهة للمكنز ، باستعمال التعبيرات مصطلح أعرض (م ع BT) ومصطلح أضيق (م ض NT) ومصطلح متصل (م ص RT) ، إلا أن العلاقات التي يتم التعبير عنها مازالت بعيدة جداً عن التحديد . ففي الجانب الأيمن للشكل رقم (٣١) على سبيل المثال ، لا تعد مصطلحات « الأطفال Children » الواردة كمصطلحات أضيق NTs تحت « الطب النفسي للطفل Child psychiatry » مصطلحات أضيق فعلاً ، لأن الأطفال ليسوا أنواعاً (فصائل) من الطب النفسي للطفل . هذا بالإضافة إلى أن دايكسترا ينبه إلى أن قائمة

رءوس موضوعات مكتبة الكونجرس ليست مكنزاً ، حتى وإن بدت شبيهة بالمكنز ، كما أنها لا يمكن أن تصلح لأغراض البحث في الفهارس المتاحة على الخط المباشر .

والمكانز بوجه عام من لغات الربط اللاحق ، بينما رءوس الموضوعات من لغات الربط المسبق في الأساس . ومن الممكن ملاحظة ذلك في أمثلة كل من قائمة رءوس موضوعات مكتبة الكونجرس وقائمة رءوس الموضوعات الطبية ، التي تشتمل على كثير من العبارات ، كما تتيح أيضاً إمكانية كشف الوثائق بمصطلحات أكثر تعقداً عن طريق التوافق بين الرءوس الرئيسية والرءوس الفرعية . ويؤدي هذا التعقد عندما يستخدم في الفهارس المطبوعة أو الفهارس البطاقية إلى نشأة الحاجة إلى إحالات مكثفة لتوفير نقاط وصول أو مداخل متعددة . وعلى ذلك ، فإن مصطلحات مثل « المقابلة في الطب النفسي للطفل INTERVIEWING IN CHILD PSYCHIATRY » يتطلب إحالة انظر CHILD PSYCHIATRY, INTERVIEWING IN SEE INTERVIEWING IN CHILD PSYCHIATRY وذلك لإثارة الانتباه لهذا المصطلح من جانب المهتم بإجراء بحث مكتمل عن وثائق الطب النفسي للطفل .

خطط التصنيف :

من الممكن أيضاً النظر إلى خطط التصنيف باعتبارها إحدى أدوات التحكم في اللغة؛ فهي من الناحية البنائية مماثلة للمكانز ، حيث ينبغي في الحقيقة لعلاقات المصطلحات الأعرض والمصطلحات الأضيق ، والتي تظهر في الترتيب الهجائي للمكنز أن تعتمد على تصنيف دقيق للأنواع المتفرعة عن الجنس . ويدعو فعلاً عدد من المواصفات والقواعد الإرشادية الخاصة بالمكانز ، بما في ذلك القواعد الإرشادية الخاصة باليونسكو *Unesco* (*Guidelines* , 1981) إلى مثل هذا الأساس الذي يمكن الاعتماد عليه . وتعد التقسيمات التفرعية لقائمة رءوس الموضوعات الطبية ، والتي توضحها القائمة (١) في الشكل رقم (٣٠) أمثلة جيدة للتنظيم المصنف للمصطلحات . ومن مثل هذا التصنيف يمكن أن يشتق البناء التفرعي (المصطلحات الأضيق في مقابل المصطلحات الأعرض) للمكنز ، وربما آلياً ، بواسطة أحد برامج الحاسب الآلي . ويتم التعبير عن المصطلحات ، في معظم خطط التصنيف ، بالطبع ، برموز هجائية رقمية (كما في المثال G2. 403.28 في الشكل رقم (٣٠)) كما يمكن استعمال هذه الرموز (أرقام التصنيف) بدلاً من

المصطلحات في عمليات الاسترجاع . ويقدم لانكستر (1972) Lancaster معالجة متعمقة لدور التصنيف في التحكم في اللغة واسترجاع المعلومات .
مظاهر التطوير في التصميم :

على الرغم من أن المبادئ الأساسية للتصميم الجيد للغات قد ظلت ثابتة نسبياً لسنوات طويلة (ويمكن ردها ، في بعض جوانبها ، إلى جهود كتر (1876) Cutter ، فقد طرحت بعض الاقتراحات الرامية إلى التطوير والتجويد . ومن بين هذه الاقتراحات ما طرحه كل من وانج وفاندندورب وإيفنز (1985) Wang, Vandendorpe and Evens ، وينطوي على إدخال المزيد من وسائل التمييز الدلالي بين المصطلحات ، بما يفوق ما تكفله علاقات المصطلحات الأعرض والمصطلحات الأضيق ، والمصطلحات المتصلة ، وإحالات « استعمل » ، و « يستعمل بدلاً من » ، الحالية . فبالإضافة إلى علاقة الجنس بالنوع أو العلاقات التفرعية (المصطلحات الأعرض والمصطلحات الأضيق) والمترادفات (استعمل ويستعمل بدلاً من) ترى المقترحات الحديثة تحليل علاقات المصطلحات المتصلة إلى عدد من الفئات الأكثر تخصيصاً ، كعلاقة الجزء بالكل (كالقرنين والبقرة مثلاً) ^(١) . وتدل التجارب التي أجريت على مجموعات المصطلحات التي تكونت بهذه العلاقات على أنها يمكن أن ترتفع بمستوى كفاءة الاسترجاع في بعض الحالات .

هذا ، وقد طرح المزيد من المقترحات المختلفة لتطوير رموس الموضوعات ، وخاصة رموس موضوعات مكتبة الكونجرس . وتشمل هذه المقترحات إعادة بناء رموس موضوعات مكتبة الكونجرس في شكل مكنز (Dykstra, 1988 a , b) ، وتغيير صيغة المصطلح ليكون غير مقلوب (مباشر) وأكثر ميلاً للربط اللاحق (Cochrane , 1986) . وهناك اقتناع عام بأنه لا غنى عن هذا النوع من التغييرات لكي تعمل رموس موضوعات مكتبة الكونجرس بشكل ملائم في بيئة الخط المباشر .

اللغة الطبيعية في استرجاع المعلومات :

هناك أربعة سبل محتملة لمعالجة اللغات التي يمكن أن تستخدم في التعبير عن ناتج التعرف على المحتوى الموضوعي للوثائق ، وإجراء عمليات البحث في نظم الاسترجاع :

(١) تسمح المواصفات الدولية لبناء المكنز ، الصادرة عن اليونسكو والمنظمة الدولية للتقييس ISO ، بالتمييز بين العلاقات التفرعية generic والعلاقات التبعية أو التجزئية partitive .

١ - التحكم في اللغة في المدخلات والمخرجات : يستخدم هذا السبيل لغة مقيدة مسبقاً كالمكنز التقليدي .

٢ - لا تحكم من أي نوع في المدخلات أو المخرجات : ويرتبط هذا السبيل بنظم الاسترجاع المعتمدة على اللغة الطبيعية .

٣ - التحكم في اللغة في المدخلات وعدم التحكم في مرحلة المخرجات : حيث يمكن لمستولي البحث استعمال كل ما يروق لهم من مصطلحات ، ويقوم الحاسب « برسم خريطة » هذه المصطلحات بمضاهاتها بالمصطلحات المقيدة للنظام أو بأي وسيلة أخرى .

٤ - لا تحكم في مرحلة المدخلات مع تحكم نسبي في مرحلة المخرجات : يستخدم هذا السبيل بالاعتماد على المكنز المخصص للبحث فقط ، ويمكن أن يسمى باللغة التي يتم التحكم فيها لاحقاً .

وقد حظي أول هذه البدائل بالتغطية التفصيلية في القسم السابق ، إلا أن هناك عدداً من الأحداث التي دفعت بعض الباحثين لاقتراح التخلي عن اللغات المقيدة (Cleverdon, 1984) لصالح نظم اللغة الطبيعية . وقد نشأت هذه الفكرة نتيجة لعدة أسباب ، منها :

- ١ - تزايد عدد الوثائق التي تحتاج إلى كشف .
- ٢ - تزايد كم نصوص اللغة الطبيعية التي يمكن البحث فيها ، وخاصة النصوص الكاملة للوثائق .
- ٣ - تكلفة وضع وصيانة واستخدام اللغة المقيدة .
- ٤ - تأخر ظهور بيانات الوثائق في المصادر الوراقية نتيجة للوقت اللازم لكشفها .
- ٥ - تكلف اللغة المقيدة أو بعدها عن اللغة الطبيعية ، والعزوف عن فرضها على مسئول البحث ، وخاصة المستفيد النهائي البسيط .

وتعتمد نظم استرجاع اللغة الطبيعية بوجه عام على العناصر الموجودة فعلاً في التسجيل الوراقية (أي أننا لسنا بحاجة لتعيين المصطلحات فكراً من جانب منتج مرصد البيانات) ، وتشمل هذه العناصر العنوان والمستخلص (الذي يعده المؤلف)

والنص الكامل في بعض الأحيان . ويتم تجهيز هذه التسجيلات الوراقية بواسطة متعهد الخط المباشر بالطريقة نفسها التي يتم بها تجهيز التسجيلات التي يتم تكثيفها بواسطة البشر ، إلا أن الوسيلة الوحيدة للاسترجاع الموضوعي هي الكلمات المستعملة فعلاً في العنوان أو المستخلص أو النص .

وقد سبق أن تناولنا مشكلات البحث في مثل هذه البيئة في هذا الفصل ، في معرض تناولنا لمبررات استعمال اللغات المقيدة . وفي نظم الاسترجاع المعتمدة على اللغة الطبيعية يتعين على مسئول البحث أن يبذل قصارى جهده في صياغة استراتيجية البحث ، معتمداً في الأساس على درايته بالمجال الموضوعي ، أو على المصطلحات التي يلتقطها من المستفيد الذي يتم إجراء عملية البحث من أجله . وبالنسبة للمستفيد النهائي البسيط ، الذي قد لا يدرك ضرورة استعمال جميع المترادفات والمصطلحات المتصلة دلاليّاً ، لا يمكن أن يكون البحث شاملاً فعلاً . إلا أن اللغة الطبيعية مقارنة باللغة المقيدة عدة مزايا واضحة ، بالإضافة إلى المزايا الاقتصادية الأساسية . ومن بين المزايا الجوهرية النهاية المفتوحة ؛ فبينما اللغة المقيدة عبارة عن قائمة نهائية يتعين انتقاء المصطلحات المتصلة بموضوعات الوثائق منها ، فإن تسجيلات الوثائق المعتمدة على اللغة الطبيعية وكذلك الطلبات أو الأسئلة ، تكفل إمكانيات لاختيار الكلمات وتوافق العبارات لا يحدها سوى حدود اللغة نفسها . هذا بالإضافة إلى أن اللغة الطبيعية أكثر تخصيصاً من اللغة المقيدة . ويمكن لجامعي اللغة المقيدة إقرار مستويات التخصيص الملائمة بفحص ما يقدم للنظام من طلبات . وينبغي أن تكون هذه العملية مستمرة نظراً لأنه من الممكن للمصطلح المخصص بما فيه الكفاية في وقت ما أن يصبح عاماً جداً فيما بعد . ويصدق ذلك بوجه خاص في المجالات الموضوعية سريعة التطور ، حيث يمكن لمصطلح في غاية الاتساع أن يكون ملائماً في البداية (كالحاسبات الآلية مثلاً) ثم يتضخم الإنتاج الفكري بعد ذلك ويتفرع إلى تخصصات دقيقة تدعو لاستعمال مصطلحات مخصصة (كالحاسبات المصغرة والحاسبات متناهية الصغر مثلاً) . واللغة الطبيعية مساهمة للتطورات دائماً ومخصصة بما فيه الكفاية ، بينما اللغات المقيدة غالباً ما تتقادم ، كما أنها غالباً ما تكون مغرقة في التعميم .

ومن بين محاولات التوفيق بين نظم اللغات المقيدة ونظم اللغة الطبيعية استعمال

التحكم إما في الكشف وإما في مرحلة البحث ، دون أي تحكم في الاتجاه الآخر . وقد خبرنا التحكم في مرحلة المدخلات دون التحكم التام في مرحلة المخرجات لبعض الوقت في قليل من النظم ، حيث يمكن لمسئول البحث استعمال مصطلحات مداخل معينة ، تتحول نتيجة للمضاهاة إلى مصطلحات النظام المقيدة . إلا أنه مما لا شك فيه أن الأمر يتطلب لغة مداخل مبدئية غاية في الضخامة لضمان التعرف على مصطلحات اللغة الطبيعية التي يستعملها مسئول البحث ، من جانب النظام . وقد اقترح كل من رادا ومل ولوتورنو وجونسون (Rada, Mill, Letourneau and Johnson (1988) وسيلة لتكوين مصطلحات الدخول الخاصة بقائمة رؤوس الموضوعات الطبية وتقييم هذه المصطلحات بشكل منهجي .

ويتمتع الحل التوفيقي الآخر ، وهو اللغة التي يتم التحكم فيها لاحقاً ، بكثير من المزايا التي تزكيه في النظم الإلكترونية لاسترجاع المعلومات . فإذا ما أحسن تطبيقه ، فإن هذا الأسلوب يجمع بين مزايا اللغة الطبيعية وكثير من مزايا اللغات المقيدة التقليدية . وعلى ذلك ، فإنه يمكن إجراء البحث على مستوى عالٍ من التخصص بناء على كلمات النص ، ولتكن «حسين» أو «عرفات» مثلاً ، كما يمكن إجراؤه بشكل أكثر تعميماً باستعمال مجموعات كلمات مكنز البحث ، كمجموعة «الأردن» أو مجموعة «فلسطين» مثلاً . وبهذه الطريقة يتحقق التخصيص إذا احتاج مسئول البحث الإفادة منه ، في الوقت نفسه الذي تتوافر فيه إمكانية البحث على أي مستوى من مستويات التعميم . أما في الأسلوب التقليدي للتحكم في اللغة ، فإن البحث عادة ما يكون مقيداً بمدى التخصيص في مصطلحات اللغة المقيدة ، ويمكن لذلك أن يعني أن يسفر البحث عن الوثائق المتصلة «بالمك حسين» حتماً عن استرجاع كل ما كشف تحت «الأردن» وكثير منه قد لا يكون متصلاً بالموضوع . وقد تناول كل من لوففر وفريدمان وشولتز Lefev- er, Freedman and Schultz (1972) إحدى طرق تطبيق التحكم اللاحق في اللغة . وما «الواجهات hedges» (مجموعات الكلمات والعبارات المتصلة بالموضوعات التي يتم البحث عنها بكثافة) والتي تتوافر للعديد من مرصدا البيانات المتاحة من خلال المرافق التجارية لاسترجاع المعلومات ، إلا تطبيق لهذه الفكرة (Sievert and Boyce, 1983) .

ويبدو أن الاتجاه السائد الآن بوجه عام هو أن نظم استرجاع المعلومات الوراقية

ينبغي أن توفر مقومات البحث باللغة الطبيعية واللغات المقيدة في الوقت نفسه. ويركز أولئك المهتمون بدراسة إمكانيات البحث الموضوعي في الفهارس المتاحة على الخط المباشر، على تصميمات النظم التي تكفل البحث بالكلمات المفتاحية (Markey, 1984 b)، ودعم تسجيلات الفهارس بعناصر من اللغة الطبيعية كالكتشافات وقوائم المحتويات (Cochrane, 1986). ولتوفير مقومات الطريقتين ما يبرره في الطبيعة المختلفة لهذه اللغات، كما يحظى بمزيد من الدعم بنتائج الدراسات التي تؤكد الطابع التكاملي لكل من اللغة الطبيعية واللغات المقيدة في الاسترجاع. فقد تبين من دراسات كل من كاتزر وزملائه (Katz et al. (1982) وتينوبير (Tenopir (1984)، وآخرين أنه من الممكن لكل طريقة أن تسترجع وثائق تتصل بالموضوع لم تسترجعها الطرق الأخرى. وتحدد دراسة حديثة لفايدل (Fidel (1992) العوامل التي يمكن أن تزكي البحث باللغات المقيدة وتلك التي تزكي البحث في نصوص اللغة الطبيعية. ومن دراسة لمئتين وإحدى وثمانين عملية بحث أجراها سبعة وأربعون باحثاً متمرساً، أمكن لفايدل التحقق من مختلف العوامل التي تؤثر في مفاضلة مسئول البحث بين المصطلحات المقيدة وكلمات النص. وقد تبين لها الاعتماد على النصوص في بعض المجالات الموضوعية أكثر من غيرها (على الرغم من أن هذا قد يكون أقل ارتباطاً بخصائص الموضوع أو لغته، منه بنوعية اللغات المقيدة المستعملة في مختلف مرادف البيانات، وخاصة مدى تخصيص هذه اللغات، وبنوعية التكشيف باللغات المقيدة).

الفصل السابع

اختيار مراصد البيانات وتقييمها

بمجرد أن يحيط اختصاصي المعلومات بحاجة المستفيد إلى المعلومات ، فإنه ينبغي أن يقرر أي مراصد البيانات يمكن البحث فيها . وكان هذا القرار منذ ثلاثين عاماً بسيطاً نسبياً ، في المكتبة الصغيرة أو مركز المعلومات الصغير على الأقل ، حيث كانت المراصد المرشحة هي خدمات التكشيف والاستخلاص القليلة نسبياً ، والتي كانت في متناول المستفيدين بشكل مباشر . والموقف اليوم مختلف تمام الاختلاف ، نظراً لأنه من الممكن حتى بالنسبة لأصغر المكتبات التعامل مع المئات من مراصد البيانات . ويدعم ذلك قدرات المكتبة في استرجاع المعلومات ، إلا أنه يضع المكتبي أمام تحديات كبيرة . ومكمن الخطر « أن يركز المكتبي على عدد ضئيل من المصادر التي قد تكون أكثر شمولاً من غيرها أو أكثر ألفة بالنسبة له ، حتى وإن لم تكن هي الأنسب دائماً بالنسبة لاحتياجات معينة إلى المعلومات . وللموضوعات البينية أو متعددة الارتباطات ، والتي يمكن التعامل معها عن طريق العديد من مراصد البيانات ، مشكلاتها الخاصة .

وليس من الضروري في جميع الأحيان أن يكون إجراء البحث في مصدر إلكتروني هو الأفضل والأصوب ، حيث يمكن للبحث في أحد المصادر المطبوعة أن يكون أقل تكلفة وأسرع في الوقت نفسه من التعامل مع أحد مراصد البيانات على الخط المباشر ، إذا كانت الحاجة تقتصر على بضع إشارات وراقية حديثة حول موضوع بسيط . إلا أنه حتى في هذا الموقف يمكن لمركز البيانات المتاح مباشرة على أسطوانات ضوئية مكتنزة أن يفضل

الكشاف المطبوع ، نظراً لأنه يمكن أن يقدم ناتجاً مطبوعاً مناسباً من الإشارات الوراقية أو المستخلصات .

وفي تعامله مع أنواع معينة من الأسئلة المرجعية السريعة ، يمكن للمكتبي أيضاً أن يقرر ما إذا كان من الممكن أن يستخدم مصدراً مطبوعاً أم مصدراً إلكترونياً . ويصف هافنر (Havener 1990) دراسة تمت فيها مقارنة المصادر المطبوعة بالمصادر المتاحة على الخط المباشر ، للإجابة عن الأسئلة المرجعية السريعة ، ويسجل أن الأسئلة المتعلقة بالحقائق البسيطة (كالبحت عن عنوانٍ مثلاً) كانت تجاب بشكل أسرع اعتماداً على المصادر المطبوعة ، أما المصادر المتاحة على الخط المباشر فقد أسفرت عن نتائج أفضل بالنسبة للأسئلة الخاصة بالموضوعات أو المفاهيم . ولما كانت أسئلة الموضوعات تتضمن الاستفسار عن إشارات وراقية لمقالات الدوريات ، أي أن البحث الحقيقي في الإنتاج الفكري ، ما لم يكن شاملاً ، فإنه لا يمكن أن يكون سؤالاً مرجعياً سريعاً فعلاً . وقد تناول أندرسون (Anderson 1989) الأسئلة المرجعية السريعة الحقيقية ، ويقدم دليلاً جازماً على أن البحث على الخط المباشر فعال من وجهة نظر التكلفة بالنسبة لأنواع كثيرة من الأسئلة التي عادة ما تتلقاها المكتبات العامة .

ويغطي هذا الفصل مختلف الأدوات المتاحة لاختصاصي المعلومات في اختيار مرصد البيانات ، ويناقش العديد من الدراسات التي تناولت عملية اختيار مرصد البيانات تفصيلاً ، كما يتناول أيضاً تقييم مرصد البيانات .

المصادر المطبوعة :

يصدر معظم متعهدي الخط المباشر فهرس بمرصد البيانات المتاحة من خلالهم بالإضافة إلى الوريقات التعريفية الخاصة بكل مرصد على حدة . وتقدم مثل هذه المصادر معلومات عامة عن المجال الموضوعي والحدود الزمنية للبيانات ، وإمكانات البحث ، والتكلفة ، وكل ما يتصل بمرصد البيانات . وعادة ما تشتمل هذه الفهارس على كشافات موضوعية ، إلا أنها تميل لأن تكون على مستوى عام جداً . وهذه المصادر مفيدة بالطبع ، ولكن في المقام الأول بالنسبة للخلفيات العامة عن مرصد البيانات المتاحة ، أو بالنسبة للبيانات المخصصة عن مرصد البيانات الذي يتقرر البحث فيه فعلاً . وليست لهذه الفهارس سوى قيمة محدودة في اختيار مرصد البيانات ، حيث

يمكن بالنسبة لأي موضوع بعينه ، أن توجه انتباه مسئول البحث نحو أبرز مرصد للبيانات، وإن كان من الممكن لمسئول البحث أن يتجه نحو ذلك المرصد على أي حال .

وهناك أيضاً الأدلة المطبوعة لمراسد البيانات (مثل *Cuadra / Elsevier Directory* [*Cuadra Associates, 1992*] *of Online Databases*) ومن مزايا هذه الأدلة أنها لا تقتصر على مراسد البيانات المتوافرة عن طريق متعهد بعينه ، كما أنها يمكن أن تقدم معلومات أكثر من تلك التي تشتمل عليها فهارس المتعهدين والورقات التعريفية ، إلا أنها لا تشتمل على كشافات تفصيلية لمحتويات مراسد البيانات، ومن ثم فإن أهميتها محدودة في مساعدة مسئول البحث في التعرف على أي مراسد البيانات يمكن الاستفادة منها بالنسبة لموضوع معين . فمن الممكن على سبيل المثال أن ترشد الباحث عن معلومات حول ميكانيكا تشقق صلب الآلات إلى مرصد بيانات *METADEx* ، وهو المرصد الذي يفوق كل ما عداه بوضوح من حيث الصلاحية لأي بحث عن المعادن ، كما يمكن أيضاً أن ترشد مسئول البحث إلى مرصد بيانات *COMPENDEX* ، إلا أنها لا يمكن أن تدل على أنه من الممكن أيضاً لكل من *CA Search* و *INSPEC* و *NTIS* (المركز القومي للمعلومات التقنية *National Technical Information Service*) أن تكون مصادر مهمة .

وتقوم بعض المكتبات بإعداد أدلتها الخاصة بمراسد البيانات التي تتيح إمكانية التعامل معها على الخط المباشر ؛ فقد تم على سبيل المثال إعداد كشاف موضوعي لأكثر من مئتين وخمسين مرصداً للبيانات، بجامعة ولاية نيويورك في ألباني (*Atkinson and Albany, 1986*) يشتمل على عدة مئات من المداخل أو نقاط الوصول. وعلى الرغم من اشتماله على عدد ضئيل من المصطلحات المخصصة ، فإن هذا الكشاف يركز أساساً على العموميات (فهناك حوالي ٢٠ مرصداً للبيانات تحت الطب ، وعشرة مراسد تحت الطاقة ، وستة مراسد تحت الانتحار) .

وكشاف (*BSO Referral Index (1985)* مخصص إلى حد ما . وهو كشاف مطبوع لمحتوى ٣٦ مرصداً للبيانات متاحة في دياالوج *DIALOG* . وهو كشاف نسبي ؛ فمصطلح « الزراعة » على سبيل المثال ، يرد في كثير من السياقات المختلفة ، كالاقتصاد والهندسة والإدارة والتسويق والإحصاء ، أما نقاط الوصول أو المداخل فهي تلك التي يقدمها نظام اليونسكو العريض للترتيب *Unesco's Broad System of Ordering* .

والمكانز المستعملة في تكثيف مراصد البيانات مفيدة جداً في بيان المجال الدقيق لكل مرصد من مراصد البيانات ، إلا أن المرء قد لا يكون على استعداد للرجوع إلى أكثر من مكنز واحد ، حتى وإن افترضنا توافرها ، لاختيار أفضل مرصد بيانات لكل عملية بحث. أضف إلى ذلك أن اللغات المقيدة لا تستخدم من جانب جميع المراصد . ويمكن لأداة تجمع لغات عدد كبير من مراصد البيانات المختلفة أن تكون أجدى بكثير . ونتناول مثل هذه الأداة بالوصف في القسم التالي .

كشافات مراصد البيانات :

لقد قام بعض المتعهدين بتجميع مراصد للبيانات ، تمثل في الواقع كشافات لمراصد البيانات المتاحة من خلالهم . وتشمل هذه المراصد DIALINDEX الذي تعدّه (خدمات معلومات دياالوج Dialog Information Services) وكروس CROSS (الذي تعدّه مؤسسة BRS Information Technologies) و DBI (الذي تعدّه خدمة أوربت للبحث Orbit Search Service)^(١) . وهذه الأدوات (التي تسمى أيضاً « بكشافات ملفات الإحالة "Cross-file indexes" أو « كشافات تردد المصطلحات term frequency indexes » أو « كشافات الخط المباشر online indexes » ببساطة) عبارة في الأساس عن ملفات قابلة للبحث ، تشتمل على جميع نقاط الوصول الواردة في جميع مراصد البيانات التي يتيحها المتعهد (كالمصطلحات الكشفية بالإضافة إلى الكلمات المفتاحية التي ترد في العناوين أو المستخلصات أو النصوص الكاملة) . ومن الممكن البحث فيها بطريقة لا تختلف كثيراً عن البحث في غيرها من مراصد البيانات ، وفقاً للتوافق المنطقية للمصطلحات .

والأهمية المحتملة لهذه النوعية من الأدوات يوضحها الشكل رقم (٢٢) الذي يبين أعلى مراصد البيانات مرتبة بالنسبة لإجراء بحث في DIALINDEX عن موضوع عطب المعادن الناتج عن القطع بالليزر . ولإجراء هذا البحث اختار مكتبي أكاديمي متمرس مرصد بيانات METADEX الذي كان أبرز اختيار نظراً لاهتمامه بالمعادن دون سواها. إلا أن نتائج البحث في DIALINDEX كانت تدل على أن مرصد بيانات ميتادكس

(١) خدمات الاسترجاع الوراق BRS أو أوربت من الشركات التابعة لمؤسسة تقنيات إنفورمو InfoPro Technologies (ماكسويل على الخط المباشر Maxwell Online سابقاً) ويضيف المترجم أن دياالوج قد بيع إلى مؤسسة نايت ردر Knight Ridder .

قد يكون أسوأ اختيار فعلاً ، لأن مرصد بيانات إنسبك INSPEC يشتمل على عدد من الوثائق التي يمكن أن تكون متصلة بالموضوع أكثر بكثير من تلك التي يشتمل عليها مرصد ميتادكس ، كما أن ستة مراصد بيانات أخرى (وبعضها جاء مفاجأة) تبدو أكثر جدوى بالنسبة لهذا الموضوع من ميتادكس . كذلك يوضح الشكل أيضاً أهمية أداة مثل DIALINDEX في تنبيه مسئول البحث إلى عدم تجاهل مراصد البيانات التي تركز على نوعيات بعينها من الوثائق ، كالتقارير التقنية (NTIS) ، وبحوث المؤتمرات (*El Engineering Meetings*) والنصوص الكاملة (*Kirk-Othmer Encyclopedia of Chemical Technology*) .

وتدل البيانات الواردة في الشكل رقم (٣٢) على أن استعمال هذا النوع من الكشافات أيسر بكثير مما هو عليه في الواقع فعلاً . فمراصد البيانات المختلفة في المقام الأول تستعمل مصطلحات مختلفة ، ومن ثم فإنه غالباً ما يكون على مسئول البحث أن يتفكر في جميع المترادفات المحتملة للمصطلح ، وهذه ليست بالمهمة الهينة ، كما أوضحنا في الفصل السادس . هذا بالإضافة إلى أنه من الممكن لإجراء بحث سطحي أن يسفر عن نتائج تختلف عن تلك التي يمكن أن يسفر عنها إجراء بحث أكثر وعياً أو أكثر تعمقاً . فمن الممكن على سبيل المثال لفكرة « العطب » أن تكون واردة في مقالة ما بشكل ضمني لا بشكل صريح ، ومن الممكن لمسئول البحث الواعي أن يرى استعمال مصطلحات إضافية « كالتبخّر » مثلاً ، والتي يمكن أن تكون متصلة بالعطب الناتج عن الليزر . وتجمّع الكشافات ، في النهاية ، مراصد البيانات التي تغطيها في فئات ، وعلى المستفيد أن يختار الفئات التي يمكن البحث فيها ، وغالباً ما تكون الفئة المتفردة هي أقوى الفئات اتصالاً بموضوع معين ، إلا أنه من الممكن لبعض الموضوعات أن تتقاطع مع عدة فئات (وفي DIALINDEX الآن أكثر من مئة فئة) . هذا بالإضافة إلى أنه ليس من الضروري لبعض مراصد البيانات التي يمكن أن تكون وثيقة الاتصال بفئة موضوعية معينة ، أن تكون أعضاء في الفئة ؛ فالفئات الخاصة بالبيئة والتلوث في DIALINDEX على سبيل المثال ، لا تشتمل على مرصد بيانات ENERGY الخاص بوزارة الطاقة ، على الرغم من أنه يمكن أن يكون مشتملاً فعلاً ، بالنسبة لبعض موضوعات التلوث ، كالأمطار الحمضية مثلاً ، على وثائق مناسبة أكثر من أي مرصد بيانات آخر . كما أن هذه الفئات

لا تشتمل على مرصد بيانات NTIS (المركز القومي للمعلومات التقنية) المهم أيضاً بالنسبة للموضوعات البيئية ، على الرغم من ورود هذا المرصد في الفئة الخاصة بالطاقة في DIALINDEX .

| الترتبة | مرصد البيانات | عدد الوثائق التي |
|---------|--|------------------|
| | تضاهي الاستراتيجية | |
| ١ | INSPEC | ٥٧ |
| ٢ | COMPENDEX | ٣٥ |
| ٣ | AEROSPACE | ٣٤ |
| ٤ | NTIS | ٢٣ |
| ٥ | WELDASEARCH | ١٩ |
| ٦ | DOE ENERGY | ١٧ |
| ٧ | SPIN | ١٣ |
| ٨ | METADEX | ١٢ |
| ٩ | <i>El Engineering Meetings</i> | ٧ |
| | <i>Kirk-Othmer Encyclopedia of Chem-</i> | |
| | <i>ical Technology</i> | ٧ |
| | أخرى | |

الشكل رقم (٣٢) : أعلى مرصد البيانات مرتبة بناء على اتصالها المحتمل بموضوع «عطب المعادن نتيجة القطع بالليزر».

ورغم ما يكتنفها من معوقات ، فإنه من المحتمل لهذه الكشافات أن تكون بالغة الأهمية في تحقيق أقصى درجات الدقة في اختيار مرصد البيانات ، كما يمكن أن تكون أدوات قوية في متناول مسئول البحث المتمرسين . وتفيد هذه الكشافات في اختيار أعلى المصادر إنتاجية بالنسبة لموضوع معين ، وكذلك في التحقق من المصادر الأوثق اتصالاً بمجال مكتبة معينة أو مركز معلومات معين (Byler and Ravenhall, 1988) .

الوسائل الآلية :

يمكن لمسئول البحث ، في المواقف المثالية ، أن يكون قادراً على إدخال وصف سردي

للحاجة إلى المعلومات . في شكل ما من أشكال أدوات اختيار مراصد البيانات ، الذي يمكن أن يقوم حينئذ بترتيب جميع مراصد البيانات المتاحة ، تبعاً لاتصالها المحتمل بالحاجة التي تم وصفها أو الإعراب عنها ، ويوفر إمكانية ربط مسئول البحث بأي مرصد للبيانات يقع الاختيار عليه . وقد أمكن إدراك الحاجة إلى مثل هذه الأدوات منذ حوالي عشرين عاماً (Lancaster, 1974 a) كما تواصلت الجهود الرامية لتطوير واجهات التعامل التي يمكن أن تشتمل على أداة من هذا النوع لبعض الوقت (انظر على سبيل المثال (Marcus and Reintjes (1977) و (Williams (1977) .

ولم نبلغ هذا الهدف بعد ، إلا أنه أمكن تطوير الأدوات التي تقترب منه ؛ فيصف تراوتمان وفون فلتنر (Trautman and Von Flittner (1989 على سبيل المثال ، «نظاماً خبيراً» يستعمل على حاسب آلي متناهي الصغر . وبإمكان النموذج الأولي الذي عرضا له ، أن يعطي جميع مراصد البيانات التي يغطيها رصيذاً رقمياً مركباً يحسب بناء على تسع خصائص أو سمات مختلفة ، تشمل نوعية الأوعية التي تحظى بالتغطية ، والمجال الموضوعي ، والمدى الزمني ، واللغة ، والمتلقي المستهدف . إلا أن هذا التقدم الظاهري خادع إلى حد ما ؛ فبالنسبة لجميع عمليات البحث تقريباً ، فإن العامل الحاسم في اختيار مراصد البيانات هو المضاهاة بين موضوع الطلب والتغطية الموضوعية لمرصد البيانات (وليس نوعية الوثائق أو لغتها أو أي خاصية أخرى) . وفي النموذج الأولي الذي وصفه تراوتمان وفلتنر ، يتم التعبير عن حدود تغطية مرصد البيانات بمصطلحات عامة (مقتطفة في الواقع من التقسيمات الفئوية الواردة في الأدلة المنشورة) لا على مستوى المصطلحات المخصصة التي تشتمل عليها كشافات مراصد البيانات ، ومن ثم فإن هذا التعبير يمكن أن يوجه الانتباه نحو مراصد البيانات البارزة ، لا نحو تلك التي تبدو أقل بروزاً . وكما أشار المؤلفان ، فإن هذه الأداة يمكن أن تساعد المبتدئ ، ولكنها لا يمكن أن تكون ذات نفع يذكر بالنسبة لمسئول البحث المتمرس .

وهناك الآن على المستوى التجاري بوابة عبور واحدة على الأقل لأكثر من متعهد واحد ، يمكن أن تقود خطى المستفيد من إمكانات الخط المباشر ، في اختيار مرصد البيانات المناسب . فنظام بوابة العبور إيزينيت EASYNET الذي طورته مؤسسة نظم تليبيز Telebase Systems Inc. ، يكفل الوصول إلى عدة مئات من مراصد البيانات

المتاحة من خلال دياالوج DIALOG ، وتقنيات إنفوبرو Infopro Technologies وغيرهما من المتعهدين (Hu, 1987, 1988; McCarthy, 1986; O'Leary, 1988; Van Brakel, 1988 a). ويتوافر نظام إيزينت في عدة نصوص مختلفة ، كل منها موجه لفئة معينة من المتلقين، ومن بينها نص يسمى إنفوماستر Info Master تعرضه شركة Western Union Telegraph Company. ويتيح نظام إيزينت عدة إمكانيات للمستفيد من خدمات المعلومات على الخط المباشر ، تشمل الاتصال بالعديد من المصادر المتاحة على الخط المباشر من خلال عملية وصل log-on واحدة ، وإجراء مطالبة واحد ، ولغة استجواب موحدة، بالإضافة إلى اختيار مرصد البيانات . وهو يعتمد على طريقة قوائم الاختيار لمساعدة المستفيد في تضيق مجال البحث وفي اختيار مرصد البيانات المناسب .

ومن الممكن توضيح طريقة استخدام قائمة الاختيار الخاصة بإيزينت / إنفوماستر EASYNET / Info Master التي تضم نوعيات الوثائق التي تغطيها مراصد البيانات ، فضلاً عن المجالات الموضوعية ، بمثال بسيط من هو (Hu (1987 التي أرادت استرجاع مقالات الدوريات التي تتناول الذكاء الاصطناعي ، حيث قدمت لها أول شاشة موضوعية فرصة الاختيار من بين ست فئات موضوعية عريضة ، اختارت منها فئة الحاسبات الآلية ، والعلوم والتقانة ، وقدمت الشاشة التالية سبعة تفرعات لهذا المجال العريض ، اختارت منها الحاسبات الآلية والهندسة والتقانة ، ويتشعب هذا القسم الفرعي أيضاً إلى سبع شعب ، اختارت من بينها الحاسبات الآلية. وقدمت الشاشة التالية ثلاث فئات لمعلومات الحاسب الآلي :

(١) الاستخدامات المنزلية أو الإدارية أو التعليمية .

(٢) البحوث والمعلومات التقنية .

(٣) الاتصالات بعيدة المدى .

واختارت هو الفئة الأولى ، وقدمت الشاشة التالية فرصة الاختيار من بين أنواع المواد التي تحظى بالتغطية ، وهي دوريات البحوث والدوريات الشعبية ، والنصوص الكاملة للدوريات ، والكتب المتخصصة في الحاسبات الآلية ، والموسوعات ، وقائمة بمراصد البيانات المتصلة بالموضوع . واختارت هو الفئة الأولى وطلب منها إدخال الموضوع

المحدد للبحث. وقد أدى المصطلح « الذكاء الاصطناعي ARTIFICIAL INTELLIGENCE أو AI عندما أدخلته هو إلى اختيار مرصد بيانات MICROCOMPUTER INDEX ، الذي تم البحث فيه ، حيث أخطر هو بوجود ٤٦١ وثيقة تلبي متطلبات البحث . وتعني طريقة قوائم الاختيار هذه على وجه التحديد أن المصطلحات التي يدخلها مسئول البحث ، بمجرد أن يتم تضيق المجال الموضوعي ، وتحديد نوعية الوثائق ، تقود المستفيد نحو أنسب مرصد للبيانات في تلك الفئة . إلا أن دراسة هو تدل على أن هذا غير صحيح ، وأنه بمجرد أن يستعمل مسئول البحث ، على سبيل المثال ، قائمة الاختيار للوصول إلى « مرصد البيانات الخاصة بمقالات الدوريات حول الاستخدامات المنزلية أو الإدارية أو التعليمية للحاسبات الآلية » فإنه من الممكن اختيار مرصد البيانات نفسه (وهو في هذه الحالة MICROCOMPUTER INDEX) أي كانت المصطلحات التي يدخلها مسئول البحث . ولقد بينت هو ، في الواقع ، أن خيطاً لا معنى له مطلقاً من الحروف ، عندما أدخل بدلاً من المصطلحات الموضوعية ، أدى إلى اختيار مرصد البيانات نفسه. ولم يكن من الواضح لها على أي أساس يتم الاختيار النهائي لمرصد البيانات ، وترى أنه من الممكن أن يكون للمصالح التجارية دورها .

ويقدم نظام إيزينت إمكانية أخرى ، وهي إيزينت سكان EASYNET SCAN ؛ فبمجرد أن يضيق مسئول البحث حدود البحث ، بقدر ما تسمح قوائم الاختيار ، فإنه يستطيع إدخال مصطلحات البحث ، ويحصل على عرض لمرصد البيانات التي تدخل في الفئة ، مع بيان عدد المرات التي ورد بها المصطلح أو المجموعة المؤتلفة من المصطلحات في كل مرصد من المراسد ، بالإضافة إلى مستوى المعلومات التي يقدمها مرصد البيانات (إشارات وراقية ، أو مستخلصات ، أو نصوص كاملة) . وتشتمل معظم العروض التي تظهر على هذا النحو ، أيضاً على «مؤشر مرصد البيانات الموصى به Recommended database indicator (RDI) » وهو عبارة عن رمز يدل على مرصد البيانات الذي توصي به مؤسسة نظم تليبيز Telebase Systems Inc. بأقصى قوة في نطاق فئة بعينها ، وربما لأنه مرصد البيانات الأساسي البارز بالنسبة لتلك الفئة (إرك ERIC للتربية ، و PAIS International للشئون العامة والدولية ، وهكذا ...) وفي سلسلة من أربع مقالات درس ماير ورويز (1990 a, b, c) ورويز وماير (1990) Ruiz and Meyer

مراسد البيانات التي اختارها المستفيدون من إيزينت عن طريق إمكانية سكان SCAN، وانتهيا إلى أن وجود مؤشر مرصد البيانات الموصى به RDI في العرض كان له أثر يفوق أي عامل آخر، بما في ذلك التكلفة، على اختيار المستفيد لمرصد البيانات؛ فقد كان المستفيدون يتأثرون بمؤشر مرصد البيانات الموصى به أكثر من تأثرهم حتى بعدد مرات ورود مصطلحات البحث في مرصد البيانات.

ويلقي التحليل الذي أجراه ماير ورويز مزيداً من الضوء على النتائج التي حققتها هو؛ فقد كان المرصد الذي يقع عليه الاختيار في دراسة هو، وبلا استثناء تقريباً هو المرصد المصحوب بمؤشر المرصد الموصى به، ويبدو إذن أن نظام إيزينت دائماً ما يختار المرصد الموصى به في نطاق الفئة أو الفئة الفرعية التي يتوصل إليها مسئول البحث، حيثما يتم تخصيص مؤشر مرصد البيانات الموصى به لمراسد البيانات التي تضمها تلك الفئة أو الفئة الفرعية.

وقد طورت خدمات معلومات ديالوج نظامها الخاص بواجهة التعامل مع مراسد البيانات التي يقع عليها الاختيار؛ فهزمة وصل ديالوج للأعمال DIALOG Business Connection تكفل إمكانية التعامل مع مراسد بيانات ديالوج المتخصصة في إدارة الأعمال، كما تكفل هزمة وصل ديالوج الطبية DIALOG Medical Connection إمكانية التعامل مع المصادر المتخصصة في الطب. وواجهات التعامل هذه مماثلة لإيزينت، حيث تعمل عن طريق استعمال قوائم الاختيار، ولكنها تختلف عنها لأن هذه الواجهات لا ترشد المستفيد إلى مرصد بيانات بعينه، وإنما ترشده إلى أسلوب البحث الذي يطبق حينئذ على مجموعة من مراسد البيانات. وكما يبين أوليري (1986) O'Leary فإن «المستفيد لا يعرف أي مراسد البيانات يتم فيه البحث فعلاً» (P. 15).

الدراسات المتصلة:

هناك كثير من المقالات التي تقارن بين مراسد البيانات المتخصصة في مختلف الموضوعات، من حيث التغطية فضلاً عن الخصائص النوعية. كذلك يقدم بعض المؤلفين توجيهات إرشادية حول اختيار مراسد البيانات في المجالات الموضوعية العريضة. ويتناول سنو (1985) Snow علوم الأحياء، وتقدم الشوكي وزملاؤها (1988) El-Shooky, et al. توجيهات إرشادية حول اختيار مراسد البيانات المناسبة

لتلبية طلبات المعلومات الصناعية ، مع العناية بوجه خاص باحتياجات الدول النامية .
ويناقش وانجر (1977) Wanger معايير اختيار مرصد البيانات ، ويصف تشامس
(1988) Chamis استخدام نظام باتي للتحويل اللغوي - Battelle's Vocabulary Switch-
ing System في اختيار مرصد البيانات ، على الرغم من عدم توافر هذه الأداة على نطاق
واسع.

لقد أجرت هو (1987, 1988) Hu دراسة ضافية لعملية اختيار مرصد البيانات ،
حيث قارنت الاختيارات الفعلية التي قام بها مسئولو بحث متمرسون بالاختيارات التي
قام بها طلبة الدراسات العليا في علم المكتبات بمساعدة إنفوماستر Info Master ،
للمجموعة نفسها المكونة من خمسين سؤالاً ، وانتهت إلى أن النتائج التي أسفر عنها
استخدام إنفوماستر تتوقف بشكل ملحوظ على القرار البشري الخاص بالمجال
الموضوعي العريض الذي ينضوي تحته السؤال (عملية الاختيار اعتماداً على قائمة
الاختيار) وأنه عندما يختار المستفيد من نظام بوابة العبور أفضل فئة موضوعية
بالنسبة للسؤال ، فإنه يمكن لإنفوماستر أن يقوم باختيار مرصد البيانات ، وكذلك
مسئول البحث المتمرس. وفي تسع وعشرين من عمليات البحث الخمسين ، كان مرصد
البيانات الذي وقع عليه اختيار إنفوماستر ، هو نفس المرصد الذي وقع عليه اختيار
مسئول البحث المتمرس ، وكان اختيار مسئول البحث المتمرس هو الأفضل في ثلاث
عشرة حالة ، بينما كان اختيار إنفوماستر هو الأفضل في ثماني حالات . وكان المعيار
المستخدم في الحكم على عمليات البحث هو عدد الوثائق التي يمكن أن تكون متصلة
بموضوع البحث ، والتي يشتمل عليها مرصد البيانات ، كما يتضح من إجراء البحث في
DIALINDEX . وقد أجري تقييم محدود لإيزينت من جانب فان براكل Van Brakel
(1980 a) الذي قام أيضاً بتجميع قائمة تفصيلية بالمعايير التي يمكن بناء عليها تقييم
نظم بوابات العبور . وكانت خبراته في خصائص اختيار مرصد البيانات مماثلة لتلك
التي سجلتها هو.

نوعية مرصد البيانات :

مازالت مناقشة الخصائص النوعية لمرصد البيانات في هذا الفصل تقتصر على ما
يتصل بتغطيتها للإنتاج الفكري في موضوعات بعينها . ويبدو من المناسب الآن النظر
في نوعية مرصد البيانات من منظور أعرض .

لا يمكن تقييم مرصد البيانات الوراقى بمعزل عن قدرته على الاستجابة لمختلف حالات الحاجة إلى المعلومات ، وإنما بناء على هذه القدرة في الأساس . ومن الممكن ، بالنسبة لحاجة بعينها إلى المعلومات ، تقييم مرصد البيانات الوراقى وفقاً لأربعة معايير أساسية:

١ - التغطية : ما مقدار ما يشتمل عليه المرصد من إنتاج فكري في الموضوع صادر في خلال فترة زمنية معينة ؟

٢ - القابلية للاسترجاع : ما مقدار ما يمكن استرجاعه من إنتاج فكري حول الموضوع ويحظى بالتغطية في مرصد البيانات ، وذلك باستعمال استراتيجيات بحث مناسبة ؟

٣ - القابلية للتنبؤ : اعتماداً على ما يشتمل عليه مرصد البيانات من معلومات ، إلى أي حد يمكن للمستفيد أن يقرر أي الوثائق يمكن أن تكون مفيدة وأيها يمكن أن تكون العكس ؟

٤ - الفورية : هل يمكن استرجاع الوثائق التي صدرت مؤخراً ، أم يؤدي التأخير في الكشف والاستخلاص إلى استرجاع وثائق تمثل نتائج البحوث القديمة لا الحديثة ؟

التغطية :

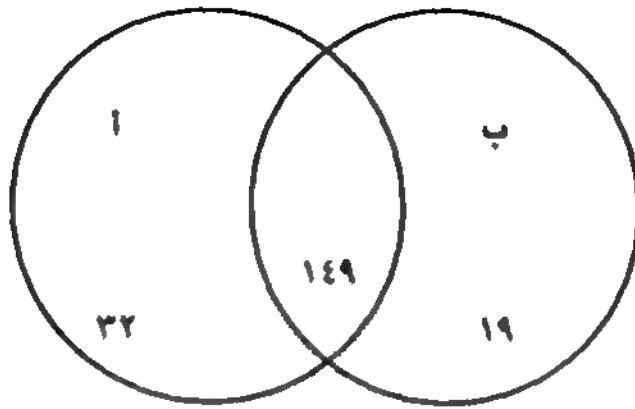
هناك تشابه بين تقييم تغطية مراصد البيانات وتقييم مدى اكتمال مقتنيات المكتبات . ومن بين طرق تقييم تغطية مقتنيات المكتبة لموضوع معين ، البحث عن وراقيات يمكن الاعتماد عليها في الموضوع ، ومراجعة هذه الوراقيات مقابل المقتنيات(*) . ومن الممكن استعمال هذا الأسلوب أيضاً في تقييم تغطية خدمات الكشف والاستخلاص . وقد شرح مارتن (1967) Martyn ومارتن وسليتر Martyn and Slater كيفية استخدام هذه الطريقة . ونفترض على سبيل المثال أننا نرغب في تقييم تغطية المدلرز/ المدلاين لموضوع الأمراض الجسد نفسية Psychosomatic ، وأنه بإمكاننا العثور على وراقية تتطلع أو تدعي أن تكون شاملة في هذا الموضوع لفترة معينة . ففي هذه الحالة يمكن أن تكون المهمة يسيرة ، حيث تراجع المداخل الواردة في الراقية

(*) تعرف هذه الطريقة بقوائم المراجعة . (المترجم) .

مقابل مرصد البيانات بناء على اسم المؤلف، لتحديد ما يشتمل عليه المرصد وما لا يشتمل عليه . ويمكن للمرء أن ينتهي نتيجة لذلك إلى أن مرصد البيانات يغطي ٨٤٪ مثلاً من هذا الإنتاج الفكري . ونحتاج بالطبع إلى الإلمام ببعض الشيء بسياسات مرصد البيانات موضوع التقييم ؛ فالمدلاين على سبيل المثال يكاد يقتصر تقريباً على مقالات الدوريات، ولا يغطي الكتب .

ويكتنف هذه الطريقة بعض مظاهر القصور ؛ فليس من اليسير العثور على الوراقيات الشاملة. هذا بالإضافة إلى أن المرء قد لا يدري شيئاً عن كيفية تجميع الوراقية؛ فإذا كانت الوراقية قد تم تجميعها معتمدة أساساً على المدلاين MEDLINE أو الكشاف الطبي *Index Medicus* (المقابل المطبوع للمدلاين تقريباً) فإنها قد لا تفيد كثيراً في تقييم مرصد البيانات هذا ، وإن كان من الممكن أن تفيد في تقييم مرصد آخر. إلا أن الأمر لا يتطلب وراقية شاملة فعلاً لتقدير التغطية بالنسبة لموضوع معين، وكل ما يمكن أن نحتاج إليه عينة ممثلة من الوثائق أو التسجيلات . ومن بين سبل الحصول على عينة ، الاعتماد على مرصد للبيانات كمصدر للمواد التي يمكن بها تقييم تغطية مرصد آخر . لنفترض على سبيل المثال ، أننا أردنا التعرف على مدى اكتمال تغطية مرصد بيانات COMPENDEX بالنسبة لموضوع المواد فائقة الموصلية *superconductors*، حينئذ يمكن الرجوع إلى مرصد بيانات إنسبك INSPEC حيث يتبين لنا وجود مئتي وثيقة مثلاً كشفت في هذا المرصد تحت المواد فائقة الموصلية أو فرط الموصلية *superconductivity* ، ثم نتخذ هذه المجموعة أساساً لتقدير تغطية مرصد بيانات COMPENDEX . وبعد مراجعة هذه المواد بناء على اسم المؤلف يمكن أن يتبين لنا وجود مئة واثنين وأربعين من بين الوثائق المئتين في COMPENDEX ، وبذلك يتم تقدير تغطيته بواحد وسبعين بالمئة. ولا أهمية هنا لحقيقة أن المئتي وثيقة ليست كل ما نشر عن المواد فائقة الموصلية ، وإنما تمثل هذه الوثائق بشكل ما ، مجموعة ممثلة للوثائق حول المواد فائقة الموصلية ، ومن ثم عينة مناسبة للاستخدام في تقدير التغطية .

وبالسهولة نفسها يمكن الاعتماد على المواد المسترجعة من مرصد بيانات COM-PENDEX في تقييم تغطية مرصد إنسبك INSPEC . وبهذه الطريقة يمكن أيضاً التعرف على مدى التداخل ومدى التفرد في تغطية مرصدين أو أكثر ، كما في المثال التالي :



ربما كان من الممكن الحصول على هذه النتائج بسحب عينة عشوائية من الوثائق التي تتناول المواد فائقة الموصلية من مرصد البيانات (أ) ومراجعتها مقابل مرصد البيانات (ب) ، وسحب عينة عشوائية من الوثائق التي تتناول المواد فائقة الموصلية من مرصد البيانات (ب) ومراجعتها مقابل مرصد البيانات (أ) . ويمكن لمثل هذه العينات أن تسمح بتقدير تغطية المرصد (أ) (١٨١ [١٤٩ + ٣٢] من ٢٠٠ ، أي حوالي ٩٠ بالمائة في هذا المثال) وتقدير تغطية المرصد (ب) (١٦٨ [١٩ + ١٤٩] من ٢٠٠ ، أي حوالي ٨٤ بالمائة) وكذلك تقدير مدى التداخل أو التكرار بين المرصدين (١٤٩ / ٢٠٠ ، أي حوالي ٧٥ بالمائة) ومدى التفرد (حوالي ٦ بالمائة من المواد التي يشتمل عليها المرصد (أ) ، أي أن ٢٠٠ / ٣٢ لا تحظى إلا بالتغطية في هذا المرصد ، بينما الرقم المقابل بالنسبة للمرصد (ب) أقل من ١٠ بالمائة بقليل [٢٠٠ / ١٩] . ومن الممكن الحصول على الأنواع نفسها من النتائج ، وبشكل أيسر إلى حد ما ، إذا ما سحبت عينة من مصدر ثالث (جـ) لتقدير التغطية والتداخل والتفرد بالنسبة لكل من (أ) و (ب) .

وإنه لأمر ممل بالطبع إدخال المئات من أسماء المؤلفين للتحقق من مدى اكتمال تغطية بعض مرصدي البيانات المسجلة على الأسطوانات الضوئية المكتنزة أو المتاحة على الخط المباشر . والحل في مثل هذه الحالة هو البدء بإجراء بحث موضوعي عريض ، ثم بعض عمليات البحث التكميلية باسم المؤلف . واعتماداً على المثال نفسه ، فإنه من الممكن بعد سحب عينة من وثائق المواد فائقة الموصلية من مرصد بيانات إنسبك INSPEC ، البحث في مرصد COMPENDEX بمصطلحات المواد فائقة الموصلية لمعرفة مقدار ما يمكن استرجاعه من وثائق العينة . ومن الممكن إجراء عمليات البحث بأسماء المؤلفين بعد ذلك لمعرفة ما إذا كانت الوثائق الأخرى من العينة قد وردت فعلاً في مرصد COMPENDEX أم لا ، وإذا كانت قد وردت فعلاً فكيف تم تكييفها ؟ .

وعند سحب عينات من مرصد للبيانات لتقييم مرصد آخر ، فإنه ينبغي مراعاة تواريخ النشر ؛ فمن الممكن على سبيل المثال سحب عينة من الوثائق التي يشتمل عليها مرصد بيانات إنسبك في عام ١٩٨٧ م ، وعند استعمال هذه العينة لتقييم تغطية COMPENDEX فإنه من المفترض مراجعة كشافات المؤلف بالنسبة لعام ١٩٨٧ م أولاً ، ثم مراجعة المواد التي لا يتم العثور عليها في هذه الكشافات مقابل عام ١٩٨٨ م (وربما بعد ذلك) ، أو اعتماداً على تاريخ النشر - مراجعة عام ١٩٨٦ م وربما قبل ذلك على أساس أنه ليس من الضروري لأي من منتجي مرصد البيانات أن يكون قد كشف الوثائق في الإطار الزمني نفسه لمنتج آخر .

ومن بين المصادر الأخرى لتقييم تغطية مرصد البيانات ، الإشارات الوراقية التي ترد في مقالات الدوريات . لنرجع إلى مثال سابق ونفترض أنه بالإمكان التعرف على عدد كبير من مقالات الدوريات التي صدرت حديثاً ، والتي تتناول الأمراض الجسد نفسية . فمن الممكن الاعتماد على الإشارات الوراقية الواردة في هذه المقالات لتكوين وراقية يمكن أن تستخدم لتقييم تغطية مرصد بيانات المدلاين MEDLINE أو مرصد بيانات المقتطفات الطبية *Excerpta Medica* لهذا الموضوع .

وهناك اختلاف واضح واحد بين الاعتماد على المواد المستقاة من وراقيات حول الأمراض الجسد نفسية (أو المواد المكشفة تحت هذا المصطلح في بعض الأدوات الوراقية) والاعتماد على الإشارات الوراقية الواردة في مقالات الدوريات ؛ فمن المفترض أن الأولى وثائق تتناول الأمراض الجسد نفسية في حد ذاتها ، بينما الثانية هي المصادر التي يحتاج إليها الباحثون العاملون في هذا المجال . ومن الممكن أن نتوقع للثانية أن تتجاوز حدود الموضوع المحدد ، كما أنها يمكن أن تشمل فعلاً مجالاً عريضاً من العلوم البيولوجية والسلوكية ، وربما بعض المجالات الأخرى . وربما يتراءى للمسئول عن التقييم استبعاد أي من المواد التي تبدو بعيدة عن بؤرة اهتمام موضوع التقييم ، أو يمكن أن يشملها بالتغطية على أساس أنه ينبغي للأداة الوراقية المفيدة بالنسبة للدارس ، في موضوع بحثه ، أن تكفل الوصول إلى جميع الوثائق المتصلة بهذا الموضوع والتي تدعو الحاجة إليها لدعم البحث في هذا المجال .

ومن الممكن عند تقييم مرصد البيانات الذي يقتصر كلية تقريباً على مقالات

الدوريات (كما هو الحال في المدلاين) أن نسلط طريقاً مختصراً للخروج بتقدير للتغطية ؛ فبعد سحب عينة من مصادر أخرى ، يمكن التعرف على مقالات الدوريات ، ثم إجراء المراجعة لمعرفة ما إذا كانت الدوريات نفسها تحظى بالتغطية بشكل منتظم في مرصد البيانات . ومن الممكن لهذه الطريقة ، في جميع الاحتمالات ، أن تسفر عن تقدير مقبول للتغطية . إلا أننا إذا أردنا أن نكون أكثر دقة ، فإنه يمكن مراجعة وثائق العينة (أو على الأقل مجموعة فرعية يتم سحبها عشوائياً) باسم المؤلف ، نظراً لأنه من الممكن تكثيف بعض الدوريات انتقائياً فقط ، ومن ثم عدم تكثيف بعض المقالات (وربما أعداد كاملة من بعض الدوريات) التي كان ينبغي تكثيفها ، لأي سبب من الأسباب^(١) . وعناوين الدوريات أقل فائدة في تقييم تغطية مرصد البيانات الذي يشتمل على مختلف أنواع المواد المنشورة ، ولا فائدة منها على الإطلاق في حالة مرصد البيانات المفرق في التخصص ، والذي يحاول تغطية كل شيء في موضوع معين ، أيّاً كان المصدر ، ولا يقتصر على مجموعة بعينها من الدوريات(*) .

وهناك عدة دوافع محتملة لتقييم التغطية ؛ فمن الممكن على سبيل المثال ، لمركز المعلومات أن يرغب في معرفة ما إذا كان هناك مرصد بيانات بعينه يغطي أحد المجالات بشكل شامل مكثف ، أو إذا كان المركز بحاجة لأن يعتمد على عدة مراصد للبيانات لتحقيق تغطية أكثر اكتمالاً . كذلك يمكن لمنتج مرصد البيانات أن يكون مهتماً بالتعرف على مدى نجاحه في تغطية موضوع معين . وربما كان من المهم في هذه الحالة معرفة أي نوعيات المطبوعات تحظى بالتغطية المناسبة وأي النوعيات ليست كذلك . ولتحقيق ذلك فإن الأمر قد يتطلب تقسيم المواد التي تحظى بالتغطية وتلك التي لا تحظى بالتغطية إلى فئات ، وفقاً لبعض الخصائص كنوعية الوثائق واللغة ومكان النشر وإسم الدورية . واعتماداً على هذه البيانات يمكن التعرف على كيفية الارتفاع بمستوى التغطية بأكثر السبل فعالية من وجهة نظر التكلفة .

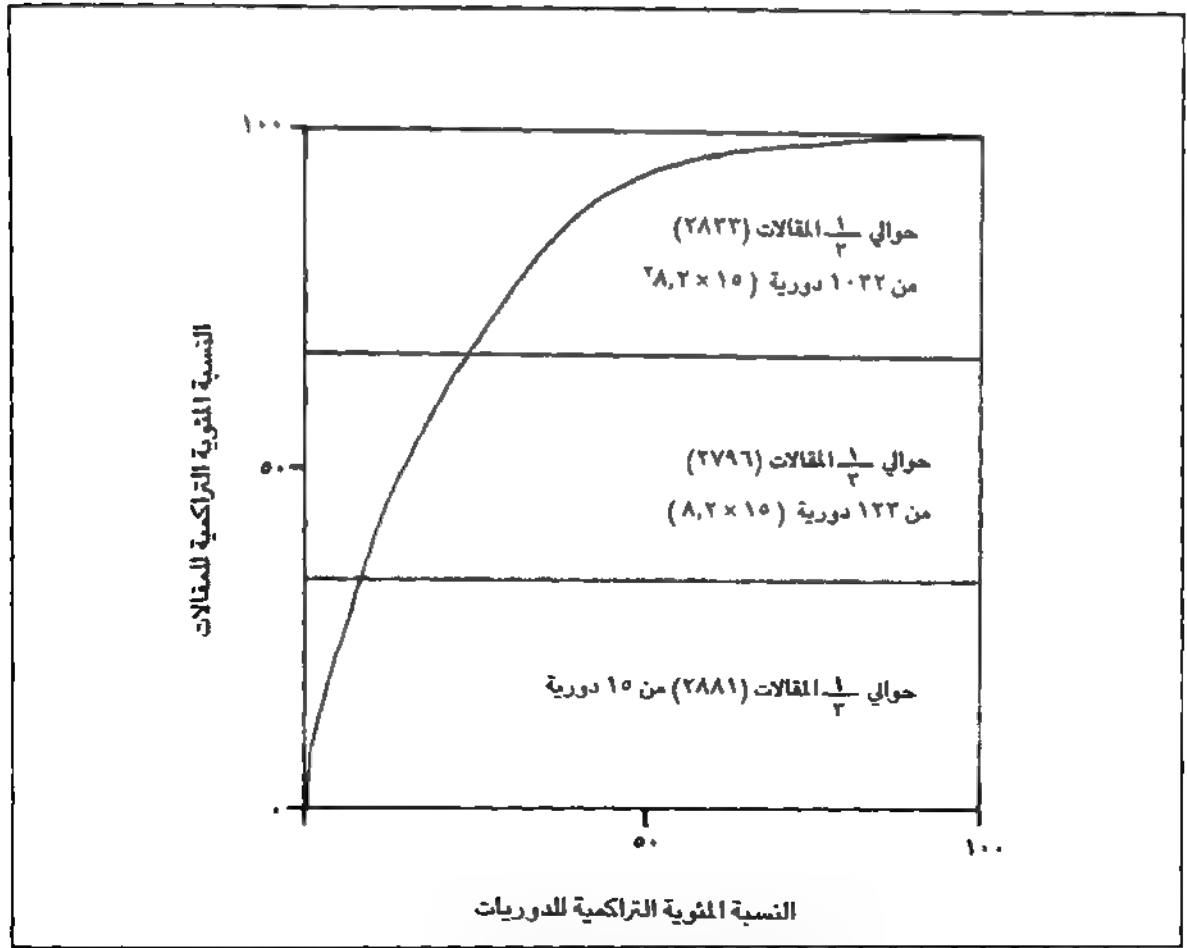
(١) لقد انتهى ثورب (Thorpe (1974) على سبيل المثال ، عند دراسته للإنتاج الفكري في مجال الروماتيزم إلى تقدير لتغطية الكشاف الطبي Index Medicus بناء على عناوين الدوريات ، مختلف إلى حد ما عن التقدير الذي انتهى إليه بناء على مقالات الدوريات . كذلك يقدم برينتن وروبرنس (Brittain and Roberts (1980) دليلاً على الحاجة إلى دراسة التغطية والتداخل على مستوى المقالات .

(*) يرجع ذلك إلى نشأت الإنتاج الفكري المتخصص في أي موضوع . (المترجم) .

ومن المهم عند النظر في تغطية مرصد البيانات الإلمام بظاهرة التششت ؛ فالتشتت يعمل في غير صالح مرصد البيانات المغرقة في التخصص ، وكذلك المكتبات ومراكز المعلومات المغرقة في التخصص ، ويميل لجانب المرصد أو المكتبات أو المراكز ذات الاهتمامات العامة . انظر على سبيل المثال في مركز معلومات متخصص في الإيدز (مرض نقص المناعة المكتسبة *acquired immunodeficiency syndrome*) الذي يريد تجميع كل الإنتاج الفكري في هذا الموضوع لإنشاء مرصد بيانات شامل . ومما يوضح أبعاد هذه المشكلة أنه لم ينشر عن الإيدز حتى نهاية عام ١٩٨٢م سوى أربع وعشرين مقالة بالدوريات ، إلا أنه بحلول عام ١٩٨٧م نما الإنتاج الفكري ليصل إلى ٨٥١٠ وثائق (Self, Filardo and Lancaster, 1989). وفي عام ١٩٨٢م كان إجمالي الإنتاج الفكري عن الإيدز موزعاً على ثلاث لغات ، إلا أنه بحلول عام ١٩٨٧م أصبح عدد اللغات المستعملة في الإنتاج ٢٥ لغة ، كما بلغ عدد الدول التي تسهم في هذا الإنتاج ٥٤ دولة . والأوقع من كل ما عداه أنه كان من الممكن تتبع إجمالي الإنتاج الفكري في موضوع الإيدز في ١٤ دورية فقط عام ١٩٨٢م ، إلا أنه بحلول عام ١٩٨٧م بلغ عدد الدوريات التي تسهم في هذا الإنتاج حوالي ١٢٠٠ دورية! ويصور هذا المثال ظاهرة التششت بوضوح ؛ فمع نمو الإنتاج الفكري لأي موضوع يزداد تششت هذا الإنتاج (مزيد من الدول واللغات والدوريات وأنواع الوثائق التي ترتبط بالموضوع) ومن ثم تزداد صعوبة التعرف عليه وتجميعه وتنظيمه .

ويرتبط أشد جوانب التششت مأساوية بانتشار المقالات في الدوريات . وكان برادفورد *Bradford* أول من لاحظ هذه الظاهرة عام ١٩٣٤م، وتعرف الآن بقانون برادفورد للتشتت. ويصور الشكل رقم (٣٣) التششت بوضوح، حيث يوقع النسبة المئوية التراكمية لمقالات الإيدز التي نشرت من عام ١٩٨٢م حتى عام ١٩٨٧م، مقابل النسبة المئوية التراكمية للدوريات التي نشرت بها المقالات . ومن الملاحظ أنه كلما ارتفعنا مع المنحنى يزداد تششت المقالات في الدوريات بمعدل هندسي تقريباً؛ فالثلث الأول من المقالات أنتجته ١٥ دورية، بينما نشر الثلث الثاني في ١٢٣ دورية (٨,٢×١٥) والثلث الأخير في ١٠٣٢ دورية (٢٨,٢×١٥) . ومثل هذا التوزيع برادفوردي تماماً . فالدورية التي تتربع على القمة في الترتيب قد نشرت في الواقع ٥٥٠ مقالة عن الإيدز خلال ست سنوات ، بينما أسهمت الثانية بـ ٣٥١ مقالة والثالثة بـ ٣٠٧ مقالات.

وكانت هذه الدوريات الثلاث وحدها مسئولة عن خمس المقالات التي نشرت خلال هذه الفترة تقريباً .



الشكل رقم (٣٣) : توقيع تشتت الإنتاج الفكري للإيدز . من ١٩٨٢ - ١٩٨٧ م.

ومما لا شك فيه أنه لا يمكن لمركز المعلومات الذي ينشئ مرصداً للبيانات في موضوع الإيدز أن يكون هذا المصدر بالاشتراك في جميع الدوريات التي تنشر مقالات في الموضوع ، وإنما يمكن للقائمة الطباقية للدوريات التي تهتم بالموضوع أن تساعد في التحقق من الدوريات البؤرية التي يمكن أن تكون جديدة بالشراء والمراجعة بشكل منتظم . ويمكن لدى قدرة المركز على تجاوز الدوريات البؤرية والنزول دونها في القائمة الطباقية ، أن تتوقف جزئياً على موارده المالية ، إلا أنه حتى ولو كانت موارده بلا حدود ، فإنه لا يمكن للمركز اقتناء جميع الدوريات التي تنشر في موضوع معين . وكلما أوغلنا نزولاً في القائمة الطباقية تضاعفت احتمالات التنبؤ بعناوين الدوريات . وعلى ذلك ، فإنه يمكن للدوريات العشر التي تحتل القمة بالنسبة لموضوع ما في المدة من عام ١٩٨٦ م

حتى عام ١٩٩٠ م ، أن تظل هي العشر التي تحتل القمة بالنسبة للسنوات الخمس التالية ، على الرغم من احتمال صدور دورية جديدة تركز أساساً على الموضوع ، وتحتل مكانها ضمن العشر التي تصدر القائمة من عام ١٩٩١ م فصاعداً . وليس من المستبعد إذن أن تظل الدوريات التي تحتل قمة الترتيب الطبقي في موضوع ما ، ضمن أعلى الدوريات إنتاجية في ذلك الموضوع لفترة ما في المستقبل . والدوريات الواقعة في منتصف الترتيب الطبقي أقل قابلية للتنبؤ ، حيث يمكن أن تواصل نشر المقالات في الموضوع أو لا تواصل . أما تلك الدوريات الواقعة أدنى الترتيب الطبقي فلا يمكن التنبؤ بها على الإطلاق ؛ فالدورية التي أسهمت بمقالة واحدة فقط في موضوع ما خلال خمس أو ست سنوات ، قد لا تسهم بأخرى على الإطلاق . وعلى الرغم من أن برادفورد قد اكتشف ظاهرة التشتت في العلوم ، كما اهتمت الغالبية العظمى من الدراسات الأخرى في التشتت بالموضوعات العلمية والتقنية ، فقد أكد ووكر وأتكينسون Walker and Atkinson (1988) وجود الظاهرة في الإنسانيات أيضاً . وعند محاولة إنشاء مرصد بيانات متخصص ، يتعين على مركز المعلومات إذن تغطية بعض الإنتاج الفكري بالاشتراك المباشر ، والتعرف على الوثائق المناسبة الأخرى بإجراء عمليات البحث المنتظمة في مرصد البيانات ذات المجالات المعرض .

وقد أجرى مارتن (Martyn (1967 ومارتن وسليتر (Martyn and Slater (1964 دراسات قيمة لتغطية خدمات الكشف والاستخلاص ، كما يتوافر العديد من الدراسات الأخرى للتغطية أو التداخل ، أيضاً في الإنتاج الفكري . ويبدو أن الكشاف الطبي Index Medicus كان أوفر حظاً من غيره في هذه الدراسات ، إلا أن وراقية الزراعة Bibliography of Agriculture كانت موضوعاً لأكثر دراسات التغطية تركيزاً ؛ ففي تقريرين متصلين ببعضهما البعض قارن بورن (Bourne (1969 a, b تغطية هذه الأداة بتغطية خمس عشرة خدمة أخرى ، وقد تغطيتها لموضوعات معينة اعتماداً على الوراقيات المصاحبة للفصول الواردة في المراجعات السنوية .

القابلية للاسترجاع :

لتغطية مرصد البيانات أهميتها بالنسبة لمن يبحث عن المعلومات حول موضوع معين ، خاصة إذا كان بحاجة إلى بحث شامل . ولقابلية الاسترجاع أهميتها أيضاً ؛ فإذا

كان مرصد البيانات يشتمل على س من الوثائق حول موضوع ما (يمكن تحديد قيمة س عن طريق دراسة التغطية) فكم من هذه الوثائق يمكن استرجاعه عند البحث في مرصد البيانات ؟

ومن الممكن اختبار القابلية للاسترجاع بدراسة تعد مكملة لدراسة التغطية . لنفترض مثلاً أننا نريد دراسة التغطية والقابلية للاسترجاع في عدد من المجالات الموضوعية الواقعة في نطاق تغطية مرصد بيانات أجريكولا AGRICOLA ، وقد أمكن بالنسبة لكل واحد من الموضوعات العشرة العثور على مجموعة من التسجيلات الوراقية ، بإحدى الطرق التي سبق أن تناولناها ، وكان من المعروف بالنسبة لكل مجموعة ما هي المواد التي يشتمل عليها أجريكولا وتلك التي لا يشتمل عليها ، فإنه من الممكن إجراء بحث لكل موضوع بواسطة اختصاصي معلومات متألف مع أجريكولا ، والحكم على القابلية للاسترجاع على أساس نسبة الوثائق المعروفة التي أمكن لمسئول البحث استرجاعها . ففي البحث الأول على سبيل المثال ، حول مخاطر الحشرات بالنسبة لفلو الصويا ، كان من المعروف أن أجريكولا يشتمل على ٨٠ وثيقة حول هذا الموضوع ، إلا أن مسئول البحث لم يتمكن من العثور إلا على ٦٠ وثيقة منها فقط ، بنسبة استدعاء ٧٥ بالمئة .

مثل هذا النوع من الدراسات يختبر ما هو أكثر من مرصد البيانات وتكشيفه ، إنه يختبر أيضاً قدرة مسئول البحث . ومن الممكن الحد من أثر هذا المتغير بأن يعهد بمهمة إجراء البحث نفسه بواسطة عدد من اختصاصيي المعلومات ، كل يعمل مستقلاً عن الآخر ، وذلك للتعرف على النتائج التي يمكن توقعها في البحث عن هذا الموضوع « في المتوسط » . ومن الممكن النظر إلى النتائج باعتبارها احتمالات ، بالإضافة إلى نسب الاستدعاء ؛ فإذا أمكن على سبيل المثال العثور على ٥ وثيقة من ٨٠ من جانب مسئول البحث الثلاثة معاً (حينئذ يكون احتمال استرجاع هذه الوثائق الخمسين / ١٠٠) وإذا أمكن العثور على ٦ من ٨٠ من قبل اثنين من مسئولَي البحث الثلاثة (بذلك يكون احتمال الاسترجاع ٠,٦٦) وإذا أمكن استرجاع ٤ من ٨٠ وثيقة من قبل واحد فقط من مسئولَي البحث الثلاثة (حينئذ يكون احتمال الاسترجاع ٠,٣٣) وإذا لم يستطع أحد العثور على ٢٠ من ٨٠ وثيقة (فإن احتمال الاسترجاع يكون صفراً) .

ولا ننسى أن القابلية للاسترجاع (الاستدعاء) لا يتم الحكم عليها إلا على أساس الوثائق المعروفة مسبقاً بأنها تتصل بموضوع البحث ، وأن مرصد البيانات يشتمل عليها فعلاً . فمن الممكن للبحث عن الآثار الضارة للحشرات على فول الصويا أن يسترجع ٢٥٠ وثيقة ، منها على سبيل المثال ١٥٠ وثيقة تبدو متصلة بموضوع البحث . وإذا أمكن استرجاع ٦٠ فقط من الثمانين وثيقة (المعروف أنها متصلة بالموضوع) فإن تقدير الاستدعاء يكون ٠,٧٥ ، بمعنى أن المئة والخمسين وثيقة المسترجعة تمثل حوالي ٧٥ بالمئة من مجموع الوثائق المتصلة بموضوع البحث في مرصد البيانات .

وترتبط نسبة الاستدعاء ببعد واحد فقط من عملية البحث . ولتقدير أو حساب التحقيق ، فإننا يمكن أن نحتاج إلى الحكم على اتصال الوثائق المسترجعة بموضوع البحث ، بطريقة ما (عن طريق مجموعة من الاختصاصيين الموضوعيين مثلاً) . ولقياس فعالية التكلفة فإننا قد نحتاج إلى معرفة التكلفة بالنسبة لكل مادة مسترجعة متصلة بالموضوع . فمن الممكن على سبيل المثال لإجمالي تكلفة بحث على الخط المباشر (بما في ذلك وقت مسئول البحث) أن يكون ٧٥ دولاراً . فإذا أمكن استرجاع ١٥٠ وثيقة متصلة بالموضوع فإن تكلفة الوثيقة المتصلة بالموضوع تكون ٠,٥٠ من الدولار .

وهناك طريقة أخرى لدراسة قابلية الوثائق للاسترجاع من مرصد البيانات ، تنطوي على نوع من المحاكاة . فإذا كنا نعرف ٨٠ وثيقة تتصل بالموضوع ص في مرصد البيانات وبإمكاننا استرجاع وطباعة تسجيلات تدل على كيفية كشف هذه الوثائق ، فإننا يمكن أن نحكي بحثاً بتسجيل عدد الوثائق القابلة للاسترجاع تحت مختلف المصطلحات والتجميعات المؤتلفة من المصطلحات . ويوضح الشكل رقم (٣٤) مثلاً على ذلك . وفي هذه الحالة ترد ٣٨ وثيقة من الثمانين التي نعرف أنها تتصل بموضوع المواد فائقة الموصلية **superconductors** تحت مصطلح المواد فائقة الموصلية ، و ١٢ وثيقة أخرى تحت مصطلح فرط الموصلية **SUPERCONDUCTIVITY** ، ولا يمكن العثور على مزيد من الوثائق تحت هذين المصطلحين ، وإنما تحت المصطلحات أ ، ب ، ج ... ي . ويمكن أن نخلص من تحليل من هذا النوع إلى أن خمسين من الوثائق الثمانين يمكن استرجاعها ببسر ، وأنه من الممكن العثور على اثنتين وستين من الثمانين بواسطة مسئول بحث حصيف ، لأن المصطلحين أ و ب إما مرتبطان ارتباطاً وثيقاً بالمواد فائقة

الموصلية وإما مرتبطان صراحة بمصطلح المواد فائقة الموصلية بواسطة إحالات في مرصد البيانات . كذلك يمكن أن نخلص إلى أن ثمانني عشرة من الوثائق الثمانين قد لا يكون من الممكن استرجاعها نظراً لأنها لا ترد إلا تحت المصطلحات التي لا ترتبط مباشرة بالمواد فائقة الموصلية (وربما كانت تمثل تطبيقات لمبدأ فرط الموصلية) .

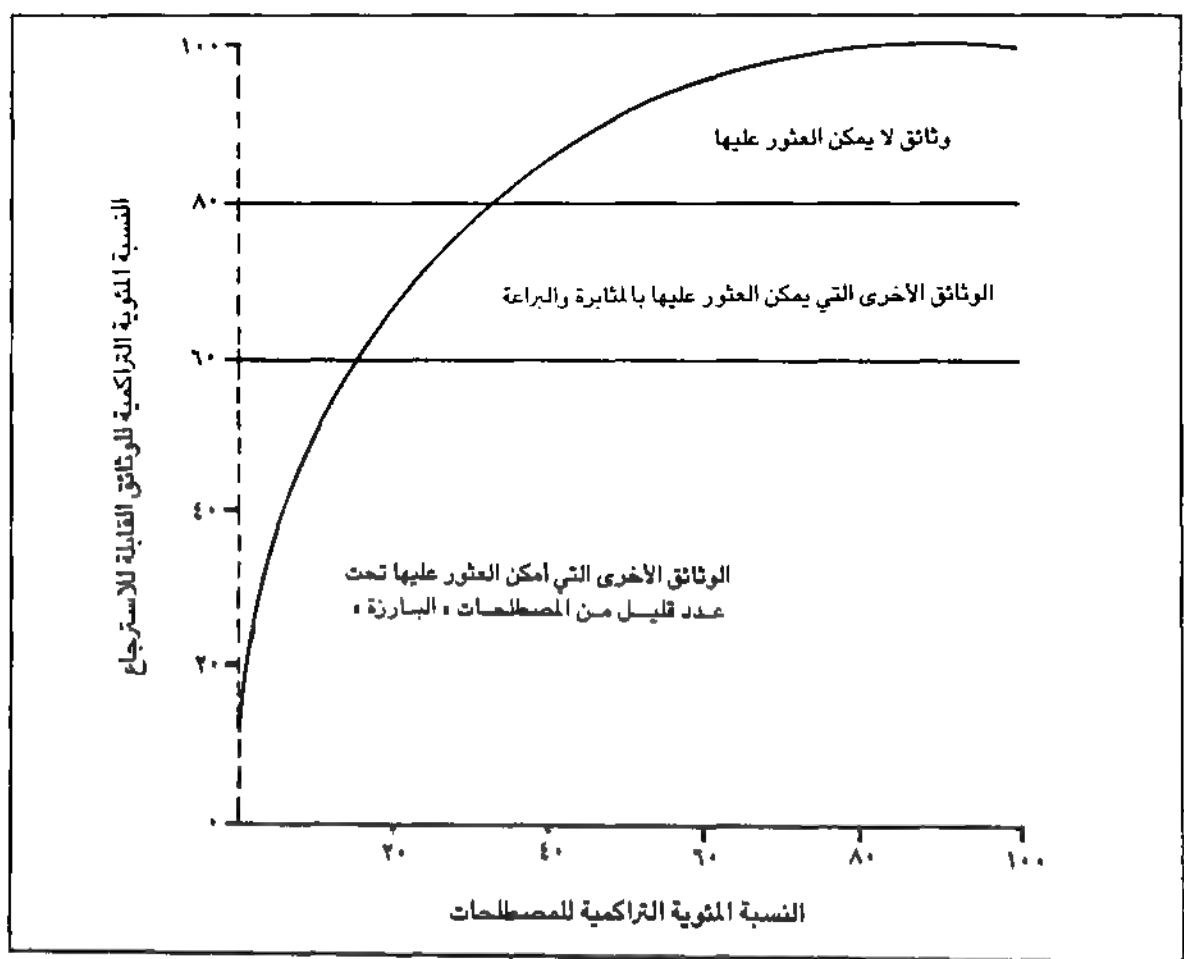
| المصطلح | عدد الوثائق القابلة للاسترجاع |
|-----------------------|-------------------------------|
| المواد فائقة الموصلية | ٣٨ |
| فرط الموصلية | ١٢ |
| أ | ٧ |
| ب | ٥ |
| ج | ٣ |
| د | ٣ |
| هـ | ٣ |
| و | ٢ |
| ز | ٢ |
| ح | ٢ |
| ط | ٢ |
| ي | ١ |
| المجموع | ٨٠ |

الشكل رقم (٣٤) : مثال لتوزيع وثائق « المواد فائقة الموصلية » تحت المصطلحات في مرصد للبيانات .

وقد أجرى البرايت (Albright (1979 دراسة تفصيلية من هذا النوع اعتماداً على الكشف الطبي *Index Medicus* ، حيث كشفت عمليات البحث التي تمت محاكاتها ، والتي أجريت لعشرة موضوعات مختلفة أن الأمر قد يتطلب ، في المتوسط مراجعة ٤٤ مصطلحاً لاسترجاع جميع الوثائق التي نعرف أنها متصلة بموضوع معين . وعلى الرغم من أن بعض هذه المصطلحات كانت ترتبط ببعضها البعض عن طريق التفرع الشجري أو الإحالات في بناء لغة مرصد البيانات ، فإن كثيراً منها لم يكن مرتبطاً على

هذا النحو ، كما لم يكن من الممكن حتى لمسئول البحث المثابر البارع البحث تحتها أو مراجعتها.

وكما تتشتت المقالات في الدوريات يمكن أن تتشتت الوثائق المتصلة بموضوع ما يضمها مرصد للبيانات على عدد كبير من المصطلحات المختلفة . ونعبر عن ذلك بيانياً في الشكل رقم (٣٥) . وربما كان من الممكن بالنسبة لأي موضوع بعينه ، العثور على نسبة مئوية عالية نسبياً ، ولتكن ٦٠ بالمئة من الوثائق المتصلة به ، تحت عدد قليل من المصطلحات البارزة « كالمواد فائقة الموصلية » أو « فرط الموصلية » عند البحث عن المواد فائقة الموصلية . وبإضافة مصطلحات أخرى وثيقة الصلة ، وربما ترتبط بهذه المصطلحات في بنية لغة مرصد البيانات ، يمكن الارتفاع بنسبة الاستدعاء إلى ٨٠ بالمئة مثلاً . ويظل هناك في هذه الحالة ٢٠ بالمئة مراوغة من الوثائق التي يشملها المرصد ، والتي قد لا يتمكن مسئول البحث من العثور عليها .



الشكل رقم (٣٥) : تشتت الوثائق تحت المصطلحات الكشفية .

وقد جاءت هذه المناقشة للمحاكاة مبسطة عن قصد ، حيث تفترض أن البحث سوف يكون مكوناً من جانب واحد ، بينما تكون معظم عمليات البحث في مرصد البيانات متعددة الجوانب أو الأوجه . وعلى ذلك فإنه عند البحث عن الآثار الضارة للحشرات على فول الصويا ، أن نفترض أن الوثيقة لا يمكن استرجاعها إلا إذا كانت مكتشفة تحت مصطلح « الحشرات » بالإضافة إلى مصطلح يدل على فول الصويا .

القابلية للتنبؤ :

لقد أكدت مناقشة تقييم القابلية للاسترجاع افتراضاً جوهرياً ، وهو أنه من الممكن التحقق من الوثيقة « المتصلة بالموضوع » اعتماداً على ما يشتمل عليه مرصد البيانات من معلومات عن الوثيقة . ويمكن لهذه المعلومات أن تشمل عنوان الوثيقة ، أو عنوان الوثيقة وقائمة المصطلحات الكشفية ، أو العنوان والمستخلص ، أو العنوان والمصطلحات والمستخلص . وكقاعدة عامة ، فإنه كلما كانت التسجيلة مطولة ، ازدادت المؤشرات التي تقدمها حول ما إذا كانت الوثيقة تهم المستفيد أم لا . وأقل ما يمكن أن يقدمه مرصد البيانات من معلومات هو عنوان الوثيقة . وكيف يعبر العنوان بالشكل المناسب على المحتوى الموضوعي ، أمر يتوقف إلى حد بعيد على نوعية الوثيقة ؛ فمقالات الدوريات المتخصصة تميل بوجه عام لأن تكون عناوينها وصفية ، بينما يمكن لمقالات الصحف استعمال العناوين الجاذبة التي قد لا تقدم وصفاً دقيقاً لمحتواها .

ولا ترد العناوين بمعزل عما عداها من عناصر بيانات الوصف ؛ ففي الكشف المطبوع ، من الممكن النظر في العنوان في إطار سياق المصطلح الكشفي الذي ورد تحته . فالعنوان « *A Rare Complication of Tuberculosis* » لا يقدم شيئاً يذكر عن محتوى إحدى المقالات ، حتى وإن كان وارداً تحت المدخل *TUBERCULOSIS, PULMONARY* (*) . أما إذا ورد العنوان تحت رأس الموضوع *AMYLOIDOSIS* (**) ، فإنه ربما كان من الممكن الحصول على فكرة أفضل عما تدور حوله المقالة . ويمكن في بعض الأحيان لاسم المجلة أو عنوان الكتاب الذي تظهر فيه المقالة أن يقدم مؤشراً للموضوع . فمقالة بعنوان « *Effects on the Presentation of Information* » لا دلالة للعنوان في حد ذاته ، أما في سياق كتاب بعنوان *Electronic Publishing* ، يصبح العنوان أكثر إحياءً أو دلالة على محتوى المقالة .

(**) التليف .

(*) السل الرئوي .

ونادراً ما يشتمل الكشاف المطبوع على قائمة كاملة بالمصطلحات الكشفية المرتبطة بالوثيقة ، إلا أنه من الممكن عادة الحصول على مثل هذه القائمة في الناتج المطبوع لمصدر البيانات المتاح على الخط المباشر ، والذي استخدم فيه الكشف البشري . ويمكن للجمع بين العنوان والمصطلحات الكشفية أن يكون غاية في القوة في الدلالة على ما تدور حوله الوثيقة . أما المستخلصات فيمكن أن تكون بالطبع ، أفضل المؤشرات دلالة على المحتوى . وإلى أي حد يمكن للمستخلصات أن تكون مؤشرات للتنبؤ هو المعيار الأساس للحكم على نوعيتها . ويقدم جينز (1991) Janes بيانات تدل على أن المستخلصات هي أكثر البدائل نفعاً للحكم على الاتصال بالموضوع ، تليها العناوين ، والمعلومات الوراقية ، وأخيراً المصطلحات الكشفية . وقد تناول لانكستر (1991) Lancaster الطرق التي يمكن بها تقييم أهمية مختلف أشكال بدائل الوثائق كمؤشرات تدل على المحتوى .

الفورية :

الفورية أو « المواكبة الآنية Currency » مقياس للسرعة التي تتم بها تغطية مرصد البيانات للوثائق الجديدة . ودراسة الفورية ليست بالمهمة اليسيرة . وعادة ما تنطوي على الخطوات الثلاث التالية :

١ - سحب عينة من الوثائق ، حوالي ٢٠٠ - ٣٠٠ في العادة ، من الملف المطبوع أو الملف الإلكتروني موضوع التقييم .

٢ - قياس المدى الزمني بين نشر الوثيقة الأولية الأصلية ونشر تسجيلاتها أو بدائلها في الملف الثانوي .

٣ - التعبير عن متوسط الفاصل الزمني بالنسبة للملف .

ويمكن الحصول على مزيد من التفصيلات حول هذه العملية في كتاب مارتن ولانكستر (1981) Martyn and Lancaster .

آفاق المستقبل :

لم تبلغ صناعة مرصد البيانات مرحلة التشبع بعد ؛ فسوف يستمر توافر مرصد البيانات الجديدة على الخط المباشر ، وخاصة في المجالات الموضوعية المتخصصة . ومما

لاشك فيه ، أنه كلما تزايد عدد مراصد البيانات ، ازدادت صعوبة التعرف عليها والمفاضلة بينها بهدف الاختيار . هذا بالإضافة إلى وجود نوع آخر من ظاهرة التششت بالإضافة إلى الأنواع التي عرضنا لها . فكما سبق أن بينا ، فإنه مع نمو الإنتاج الفكري لموضوع ما ، يتزايد تششت هذا الإنتاج على الأنواع المختلفة للوثائق ، واللغات ، والدوريات ، كما يزداد تششته أيضاً في مراصد البيانات ، وهذه ظاهرة الملح إليها بار (1988) Bar والشوكي وزملاؤها (1988) El-Shookey & al. وقد صور لانكستر ولي (1985) Lancaster and Lee هذا الشكل من التششت بوضوح ، اعتماداً على مثال الأمطار الحمضية . ففي عام ١٩٧١م ورد الإنتاج الفكري الكامل للأمطار الحمضية (وثيقة واحدة !) في مرصد بيانات واحد . ثم بدأ مرصدان آخران يهتمان بالموضوع عام ١٩٧٢م ، وبحلول عام ١٩٧٤م أصبح الإنتاج الفكري في الأمطار الحمضية موزعاً على ستة مراصد للبيانات على الأقل . ويوضح الشكل رقم (٣٦) كيف انتشر الإنتاج الفكري كما يعبر عنه عدد مراصد البيانات التي ورد بها مصطلح « الأمطار الحمضية » بحلول عام ١٩٨٣م ؛ ففي ذلك الوقت كان هناك ٤٢ مرصداً للبيانات تهتم بالموضوع . والجدير بالملاحظة أيضاً كثرة السياقات المختلفة التي حظي فيها موضوع الأمطار الحمضية بالاهتمام عام ١٩٨٣م ؛ فقد حظي بالاهتمام في صناعة الطاقة ، وفي الاهتمامات البيئية بوجه عام ، وفي الآثار البيولوجية والطبية ، وفي الكيمياء (بما في ذلك كيمياء طبقات الجو العليا) ، وفي الآثار الزراعية ، وفي الأحياء المائية ، وفي الحلول الصناعية ، وفي التأثير على المواد ... وهكذا . ومن المحتمل أنه إذا ما أجري البحث اليوم فسوف يكشف عن أعداد متزايدة من مراصد البيانات التي تهتم بالموضوع ، نظراً لأنه قد أمكن رصد تأثيرات الأمطار الحمضية في سياقات أخرى ، كتأثيراتها على سبيل المثال ، في المباني ذات الأهمية التاريخية ، وفي المواقع ذات القيمة الأثرية . كذلك يوضح الشكل رقم (٣٦) كيف يمكن لموضوع يحظى بالاهتمام أن يتوزع في مراصد البيانات التي تغطي مختلف أنواع الوثائق ، كالدوريات العلمية ، والتقارير التقنية ، والمجلات الشعبية ، وبحوث المؤتمرات ، والأطروحات ، والمطبوعات الحكومية ، والمضابط الخاصة بالمجالس النيابية ، وما شابه ذلك .

كذلك يتضح من الشكل أن مشكلة اختيار مراصد البيانات مشكلة معقدة (حيث يمكن لبحث شامل فعلاً عن الأمطار الحمضية أن ينطوي على التعامل مع العديد من

| الترتبة | مرصد البيانات | عدد مرات الورود |
|---------|---|--------------------|
| 1 | DOE ENERGY | 1,175 |
| 2 | ENVIROLINE | 729 |
| 3 | BIOSIS | 432 |
| 4 | CHEMICAL ABSTRACTS | 361 |
| 5 | NTIS | 313 |
| 6 | NATIONAL NEWSPAPER INDEX | 216 |
| 7 | AGRICOLA | 207 |
| 8 | SCISEARCH | 206 |
| 9 | CONGRESSIONAL INFORMATION SERVICE | 187 |
| 10 | MAGAZINE INDEX | 186 |
| 11 | ENVIRONMENTAL BIBLIOGRAPHY | 184 |
| 12 | CAB ABSTRACTS | 144 |
| 13 | TRADE AND INDUSTRY INDEX | 133 |
| 14 | POLLUTION ABSTRACTS | 177 |
| 15 | LIFE SCIENCES COLLECTION | 116 |
| 16 | COMPENDEX | 89 |
| 17 | EI ENGINEERING MEETINGS | 89 |
| 18 | UPI NEWS | 88 |
| 19 | CHEMICAL INDUSTRY NEWS | 86 |
| 20 | EXCERPTA MEDICA | 84 |
| 21 | NEWSEARCH | 66 |
| 22 | INSPEC | 66 |
| 23 | AQUATIC SCIENCE AND FISHERIES ABSTRACTS | 61 |
| 24 | CONFERENCE PAPERS INDEX | 58 |
| 25 | CRIS | 54 |
| 26 | AQUALINE | 48 |
| 27 | PAIS INTERNATIONAL | 48 |
| 28 | APTIC | 41 |
| 29 | GPO MONTHLY CATALOG | 38 |
| 30 | AMERICAN STATISTICS INDEX | 27 |
| 31 | BHRA FLUID ENGINEERING ABSTRACTS | 6 |
| 32 | FEDERAL REGISTER | 6 |
| 33 | GPO PUBLICATIONS REFERENCE | 6 |
| 34 | ZOOLOGICAL RECORD | 4 |
| 35 | ISMEC | 4 |
| 36 | COMPREHENSIVE DISSERTATION ABSTRACTS | 4 |
| 37 | IRIS | 4 |
| 38 | OCEANIC ABSTRACTS | 3 |
| 39 | METADEx | 3 |
| 40 | TELEGEN | 1 |
| 41 | STANDARD & POORS DAILY NEWS | 1 |
| 42 | WORLD AFFAIRS REPORT | 1 |

الشكل رقم (٣٦) : ترتيب مرصد البيانات طبقاً تبعاً لعدد مرات ورود مصطلح «الأمطار الحمضية» في ١٤ سبتمبر ١٩٨٣ م.

المصادر، كما يمكن لأفضل المصادر للبحث عن جانب بعينه من الموضوع ألا يكون هو أول مصدر يتبادر إلى الذهن) ، ويمكن لهذه المشكلة أن تزداد سوء في المستقبل . وإذا كان الأمر كذلك فعلاً ، فإنه ربما كان من المحتمل تطوير أدوات للاختيار أكثر تقدماً . وقد عبر العديد من المؤلفين عن الحاجة إلى أدوات تتضمن بعض عناصر الذكاء الاصطناعي والنظم الخبيرة ، كما تناقش الشوكي وزملاؤها (El - Shookey et al, (1988 وهو (Hu (1987 وموريس وتسنج ونيوهام (Morris, Tseng and Newham (1988 ما يمكن أن تكون عليه هذه العناصر . وتتضمن الأداة التي تناولها تراوتمان وفون فلتنر (Trautman and Von Flittner (1989 بالوصف بعض عناصر النظم الخبيرة ، كما يمكن النظر إلى إنفوماستر باعتباره يتمتع ببعض الخصائص الأولية من هذا النوع ، إلا أنه لم تظهر حتى الآن أداة لاختيار مرادف البيانات ، يمكن النظر إليها باعتبارها تعتمد على ذكاء اصطناعي حقيقي ، على الرغم من أن هو (Hu (1987 تخلع مثل هذا التقدم على أسلوب التفرع البسيط المعتمد على قوائم الاختيار والمسمى بإنفوماستر INFOMASTER .

البحث في مرصد البيانات

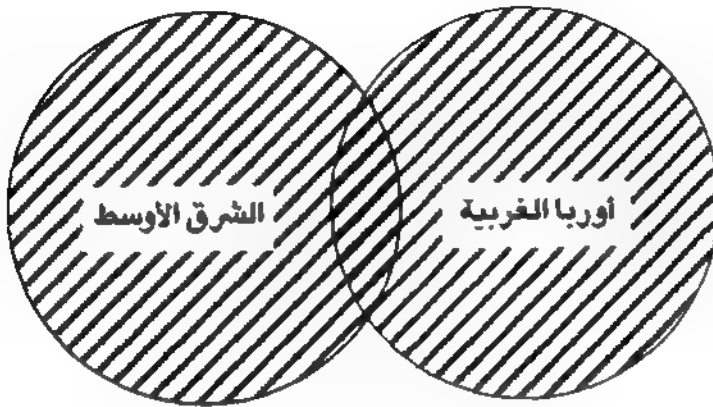
هناك تشابه كبير بين صياغة استراتيجيات البحث في مرصد البيانات وتكشيف الوثائق ؛ فمن الممكن تقسيم عملية التكشيف إلى خطوتين ؛ أولاهما التحليل الموضوعي (تحديد ما تدور حوله الوثيقة ، ودواعي اهتمام المستفيدين من خدمة معينة بها) ، وثانيتهما هي ترجمة ناتج هذا التحليل الموضوعي إلى مجموعة معينة من المصطلحات . وكذلك الحال أيضاً ، تنطوي صياغة استراتيجيات البحث على التحليل الموضوعي والترجمة ؛ فمستئول البحث يحاول أولاً التعرف على ما يبحث عنه المستفيد فعلاً ، وأي نوعيات الوثائق التي يشتمل عليها مرصد البيانات يمكن أن تفيد في تلبية الحاجة إلى المعلومات . ثم يقوم مستئول البحث بعد ذلك بترجمة ناتج هذا التحليل إلى المصطلحات التي تلائم مرصد البيانات الذي وقع عليه الاختيار . ويمكن لهذه المهمة أن تعني مجرد تحديد أي رؤوس الموضوعات يمكن البحث تحتها في الكشاف المطبوع أو فهرس المكتبة ، كما يمكن أن تعني تحديد أي المصطلحات ، وبأي تجميعات مؤتلفة ، يمكن أن تستعمل في البحث في المرصد الإلكتروني للبيانات . ويمكن الاستطراد في بيان التناظر بين التكشيف والبحث ؛ ففي التكشيف لا يمكن للتحليل الموضوعي أن يتقيد باللغة التي يمكن استعمالها ، وإنما تأتي هذه القيود فيما بعد عندما يحاول المكشف ترجمة ناتج التحليل الموضوعي إلى مصطلحات النظام (اعتماداً على أحد المكانز مثلاً) . وفي عملية البحث أيضاً تأتي القيود في مرحلة الترجمة ؛ فمستئول البحث قد لا يكون مقيداً بمجموعة

معينة من المصطلحات فحسب ، وإنما يمكن أن يواجه أيضاً قيوداً في كيفية الربط بين هذه المصطلحات في مجموعات مؤتلفة .

منطق البحث :

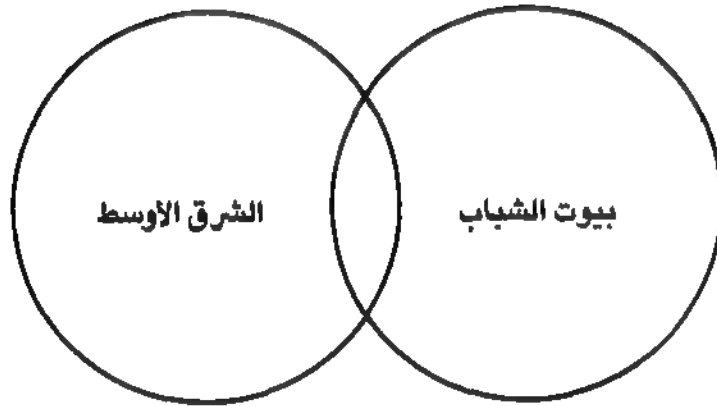
ينطوي استرجاع المعلومات على التصنيف بأوسع معانيه؛ فعملية التكشيف تنطوي في الأساس على توزيع الوثائق على فئات أو أقسام معينة ، والفئة هي المجموعة التي تضم جميع الوثائق التي خصص لها مصطلح كشفي معين (وعادة ما تنتمي معظم الوثائق إلى عدة فئات) . ومن الممكن النظر في المصطلحات الكشفية باعتبارها وسميات تميز مختلف الفئات . وينطوي البحث في مرصد البيانات على تحديد أي الفئات يمكن الرجوع إليها لتلبية الحاجة إلى المعلومات ، ومن ثم يمكن النظر إلى استراتيجية البحث باعتبارها بياناً بالفئات التي ينبغي أن تنتمي إليها الوثيقة لكي تكون متصلة بموضوع الحاجة إلى المعلومات .

ومن الممكن لاستراتيجية بحث بسيطة جداً أن تتضمن الرجوع إلى فئة واحدة فقط ، كجميع الوثائق التي تتناول الشرق الأوسط على سبيل المثال ، أو يمكن أن تتضمن فئتين يمكن لأي منهما أن تكون مقبولة ، كما هو الحال مثلاً في : الشرق الأوسط أو أوروبا الغربية ، ويمكن التعبير عنها بيانياً على النحو التالي :



ولعلاقة البديل « أو » عدة أسماء محتملة : فهي غالباً ما تسمى بالمجموع أو المجموع المنطقي للفئتين ، كما يمكن أن تسمى أيضاً بالعضوية المشتركة ، أو التناوب ، أو التناوب المنطقي ، أو الفصل ، أو التصاق الفئتين أو اتحادهما . ومن الجدير بالملاحظة أن علاقة « أو » تشمل علاقة « و » : « فالشرق الأوسط أو أوروبا الغربية »

مثلاً تعني « الشرق الأوسط أو أوروبا الغربية أو كليهما » . فالوثائق التي تنتمي لكلا الفئتين يمكن قبولها كما تقبل الوثائق التي تنتمي إلى أي منهما .

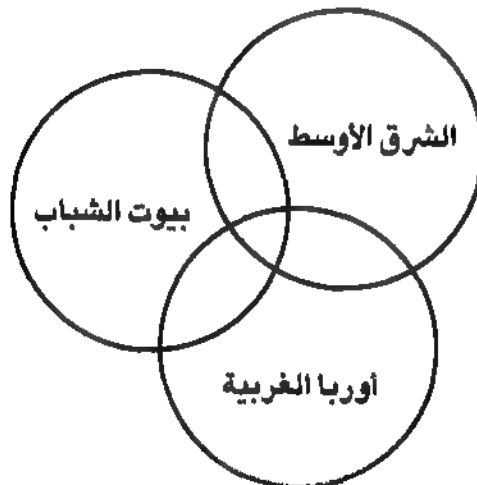


والعلاقة الرئيسة الأخرى التي تستعمل في البحث هي علاقة « و » كما في المثال الخاص « بالشرق الأوسط وبيوت الشباب » . ومن الممكن توضيح هذه العلاقة بالإشارة إلى أن المساحة المظلة تمثل المجموعة المنطقية التي يتم البحث عنها ؛ فمستول البحث يبحث عن الوثائق التي تنتمي إلى كل من فئة « الشرق الأوسط » وفئة « بيوت الشباب » . ولعلاقة « و » عدة أسماء أيضاً ، وهي العضوية الاعتيادية ، أو الاقتران ، أو الالتقاء ، أو التقاطع ، أو الناتج ، أو الناتج المنطقي للفئتين .

ولا تقتصر الحاجة إلى المعلومات ، بالطبع ، دائماً على العلاقات الثنائية البسيطة التي صورناها بيانياً دون غيرها . فلننظر في عبارة « بيوت الشباب في الشرق الأوسط أو أوروبا الغربية » ، فكمعادلة يمكن التعبير عنها على النحو التالي :

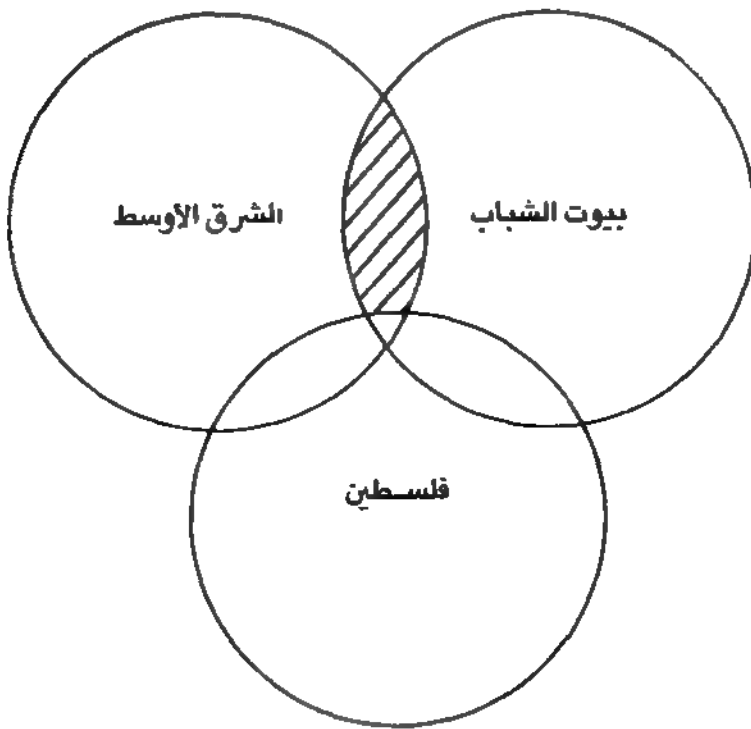
بيوت الشباب و (الشرق الأوسط أو أوروبا الغربية)

وفي شكل بياني على النحو التالي :

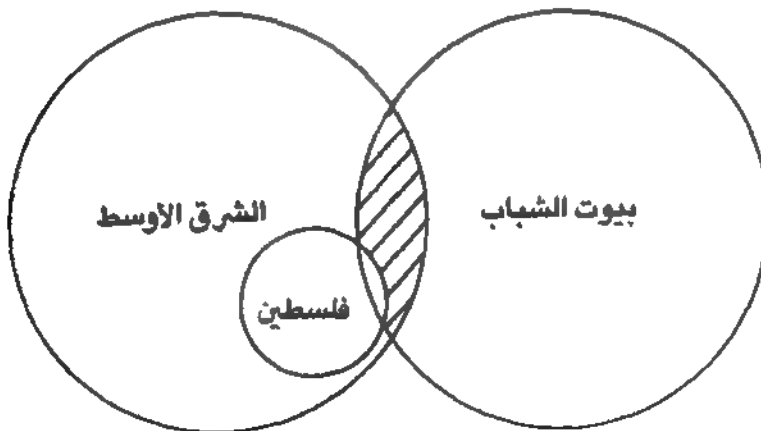


حيث تمثل أيضاً المساحة المظلة مجموعة الوثائق التي تنطبق عليها الشروط المنطقية لمتطلبات البحث .

والعلاقة الثالثة المهمة هي علاقة « فيما عدا Not » : فمن الممكن على سبيل المثال للمستفيد من مرفق المعلومات أن يعرب عن اهتمامه بالوثائق التي تتناول بيوت الشباب في دول الشرق الأوسط فيما عدا فلسطين . وعادة ما يتم التعبير عن هذه العلاقة على النحو التالي :



إلا أنه من الممكن التعبير عنها بشكل أكثر دقة على النحو التالي :



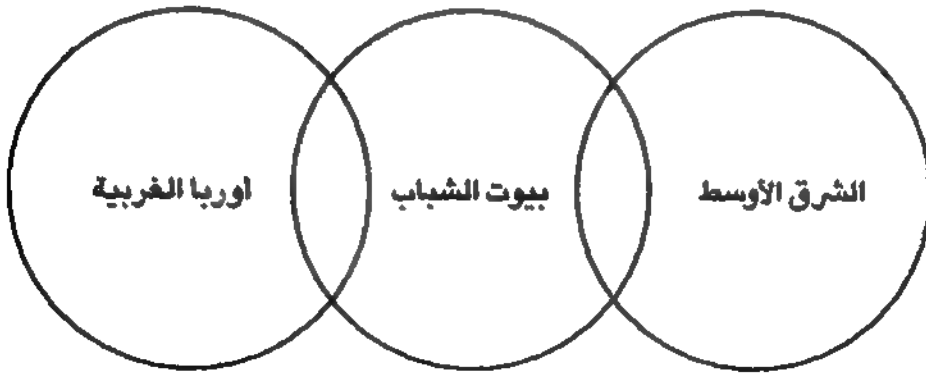
نظراً لأن فلسطين جزء من الشرق الأوسط .

وغالباً ما تسمى علاقة « فيما عدا » المنطقية بالاستبعاد (وفي هذا المثال تم استبعاد فئة فلسطين) وهي تعني التتمة المنطقية (المجموعة المنطقية التامة لفئة فلسطين). وعلى الرغم من أن الاستبعاد يمكن أن يكون مفيداً في البحث في مراصد البيانات ، فإنه يمكن أيضاً أن يساء استخدامه . وحيثما كان ممكناً ، فإن الأسلوب الإيجابي في البحث أفضل من الأسلوب السلبي . فمن الممكن ، على سبيل المثال ، للوثائق التي تتناول بيوت الشباب في عدد من دول الشرق الأوسط ، بما في ذلك فلسطين أن تكون مهمة بالنسبة للمستفيد . ومن الممكن لمثل هذه الوثائق ألا تسترجع في عملية البحث التي تنص على استبعاد فلسطين ، نظراً لأن علاقة الاستبعاد عادة ما تكون لها الأسبقية المنطقية على كل ما عداها ، في البرامج المستخدمة للبحث في مراصد البيانات ، ومن الممكن للوثائق أن تكون مكشوفة تحت فلسطين بالإضافة إلى المصطلحات الخاصة بدول الشرق الأوسط الأخرى . ومن ثم فإنه ربما كان من الممكن الحصول على نتائج أفضل بالجمع بين «بيوت الشباب» والمصطلحات الدالة على بقية دول الشرق الأوسط ، حيث لا يؤدي ذلك إلى استبعاد الوثائق التي تتناول كلاً من فلسطين وغيرها من الدول العربية في الوقت نفسه .

ويحكم العلاقات المنطقية الصورية (و ، و أو ، وفيما عدا) بين الفئات والعمليات التي تتم على هذه الفئات ، منطق متشابه ، وخاصة المنطق البولييني ، في كثير من التطبيقات المهمة . أما الأشكال التي تستخدم للتعبير عن هذه العلاقات فتسمى عادة بأشكال فن Venn diagrams ، وكثيراً ما يسيء الطلبة وبعض الكتب الدراسية ، للأسف ، استخدام أشكال فن . فهذه الأشكال ينبغي أن تستخدم للتعبير عن استراتيجيات البحث على مستوى الموضوع أو المفهوم فقط لا على مستوى المصطلح . وعلى ذلك فإن فئة «الشرق الأوسط» ليس من الضروري أن تكون مطابقة تماماً لمصطلح « الشرق الأوسط»؛ فمن الممكن على سبيل المثال أن يسترجع المصطلح تلك الوثائق التي تتناول المنطقة بوجه عام فقط ، ولا يسترجع تلك التي تتناول كل دولة على حدة . وللتعبير عن فئة الشرق الأوسط ، فإننا قد نحتاج إلى مجموعة كاملة من المصطلحات تربط بينها علاقة « أو » ، على النحو التالي :

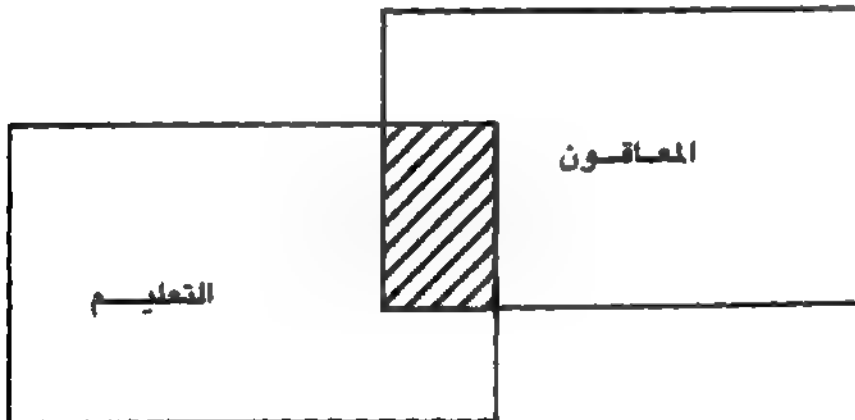
الشرق الأوسط أو الأردن أو فلسطين أو لبنان (وهكذا) .

وهناك خطأ آخر ينطوي على رسم أشكال فن التي لا تتقاطع فيها جميع الدوائر؛ ففي أشكال فن ينبغي أن تتقاطع كل دائرة مع جميع الدوائر الأخرى . والشكل التالي لا يمكن قبوله لأنه يعني استحالة وجود علاقة منطقية بين الشرق الأوسط وأوروبا الغربية:

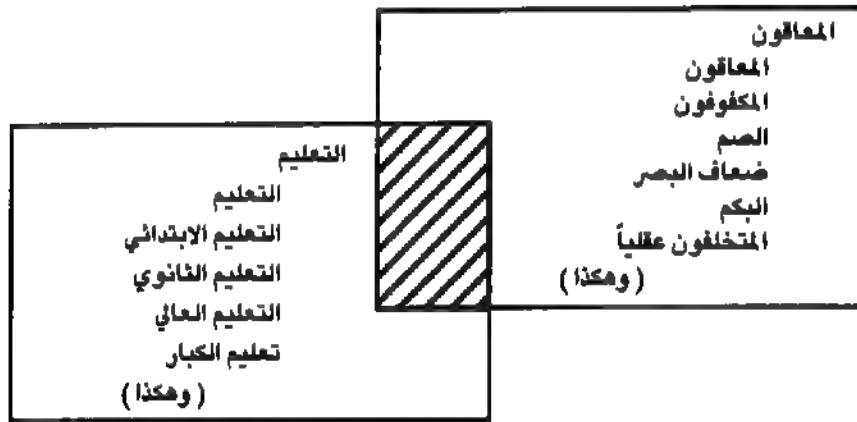


استراتيجية البحث :

تنطوي استراتيجية البحث على التحليل الموضوعي للحاجة إلى المعلومات ، وترجمة ناتج هذا التحليل إلى مجموعة من المصطلحات . وفي التحليل الموضوعي تجزأ الحاجة إلى المعلومات إلى عناصرها الفرعية أو أوجهها . فإذا كان لدينا على سبيل المثال طلب لإجراء البحث عن تعليم المعاقين ، فإن هذا الطلب يتكون من جانبين أو وجهين : « التعليم » و «المعاقون» ، وكلاهما تدعو الحاجة إليه للوفاء بمنطق حاجة المستفيد . ومن الممكن التعبير عن هذه العلاقة على النحو التالي (ينبغي أن تبين الأشكال المنطقية المساحات المتقاطعة ، إلا أنها ليس من الضروري أن تكون دوائر) :



وعند ترجمة ناتج التحليل الموضوعي إلى لغة مرصد بيانات معين ، فإن الأمر قد يتطلب عدة مصطلحات لتغطية كل جانب من الجوانب :



وقد يحتاج مسئول البحث إلى استعمال مجموعة من المصطلحات (تربط بينها علاقة « أو ») لتغطية التعليم ومجموعة أخرى من المصطلحات (تربط بينها علاقة «أو») لتغطية المعاقين ، ثم يربط بعد ذلك هاتين المجموعتين بعلاقة « و » المنطقية . وبالنسبة لعمليات البحث في مرصد البيانات الإلكترونية ، فإن المهمة ربما تنطوي على إدخال المجموعة الأولى من المصطلحات ، ثم المجموعة الثانية ، ثم يطلب من برنامج البحث الربط بين المجموعتين على النحو التالي :

١ - التعليم أو التعليم الابتدائي أو التعليم الثانوي أو التعليم العالي أو تعليم الكبار .
(٢٠٩٥)

٢ - المعاقون أو المكفوفون أو الصم أو ضعاف البصر أو البكم أو المتخلفون عقلياً .
(٨٢٨)

٢ - ١ و ٢

(٧١)

وتدل الأرقام الواردة بين الأقواس في نهاية كل جانب من جوانب البحث على أن مرصد البيانات يشتمل على ٢٠٩٥ وثيقة مكشوفة تحت مصطلح أو آخر من مصطلحات التعليم ، و ٨٢٨ وثيقة مكشوفة تحت مصطلح أو آخر من مصطلحات المعاقين ، و ٧١ وثيقة مكشوفة تحت مصطلح واحد على الأقل من كل مجموعة .

ولا يمكن لجميع الطلبات أن تكون بهذه البساطة ؛ فمن الممكن لبعض الطلبات أن تتضمن ثلاثة أوجه أو جوانب ، وربما أكثر من ثلاثة في بعض الأحيان . إلا أنه مهما بلغت الحاجة إلى المعلومات من التعقد فإنه من الممكن دائماً تقسيمها أو تحليلها إلى العناصر التي تتكون منها . وتنطوي عملية البحث في مرصد البيانات على الخط المباشر إذن على :

١ - التحليل الموضوعي .

٢ - ترجمة ناتج التحليل الموضوعي إلى مصطلحات مقبولة .

٣ - إدخال المصطلحات لتكوين مجموعات من الوثائق (فئات) .

٤ - الربط بين الفئات بطريقة تلبي المتطلبات المنطقية للحاجة إلى المعلومات .

ولا يمكن أن تكون هناك دائماً علاقة تطابق تام بين ناتج التحليل الموضوعي واللغة المستعملة في مرصد بيانات معين . فإذا نظرنا إلى طلب للمعلومات حول تأثير درجات الحرارة المرتفعة للمياه الناتجة عن الصرف الصناعي على الأسماك ، فإن مسئولى البحث يمكن أن يحلل هذا الطلب إلى ثلاثة أوجه : الأسماك ، وتلوث الماء ، والحرارة . إلا أن المكنز المستخدم في مرصد البيانات قد يشتمل على مصطلح « التلوث الحراري » الذي يعني الحرارة ، والتلوث ، والماء (عادة) . ومن ثم فإن « التلوث الحراري » مرتبطاً بمصطلحات الأسماك ، قد يكون هو المدخل الأول المفضل لهذا الموضوع ، وربما يكون من قبيل التكرار أو الحشو استعمال مصطلح « تلوث الماء » أيضاً .

ويمكن في بعض الأحيان أن تتطلب عملية بحث معينة ، منطقياً ، الربط بين أعداد مختلفة من المصطلحات معاً في استراتيجيات فرعية ؛ فإذا نظرنا إلى طلب للمعلومات حول الإنفاق على البحوث المتصلة بالقضايا العسكرية ، فإنه من الممكن البحث عن هذا الموضوع باستعمال مصطلح « الإنفاق » ومصطلح « البحث » وأية مصطلحات أخرى تدعو الحاجة إليها لتغطية الجانب العسكري . ونفترض الآن أن مصطلح « بحوث الأسلحة » موجود في مكنز مرصد البيانات ، ويعني هذا المصطلح العسكرية ، ولهذا فإنه قد يكون من قبيل الحشو ربطه بغيره من المصطلحات العسكرية . وربما أمكن للاستراتيجية المقبولة منطقياً أن تكون :

البحث و (الإنفاق أو التمويل) و (القوات المسلحة أو الأسطول أو القوات الجوية أو الجيش أو الدفاع) .

وعند ترجمة ناتج التحليل الموضوعي إلى لغة مرصد بيانات معين ، يتعين على مسئول البحث التنقيب عن المصطلحات الأكثر ملاءمة للطلب منطقياً . ويحدث في كثير من الأحيان ألا يكون بيان الطلب قابلاً للترجمة إلى اللغة المقيدة ، بدقة . فإذا كان مصطلح « تلوث الماء » على سبيل المثال ، موجوداً في مرصد البيانات ، فإنه يتعين البحث عن موضوع « تلوث البحيرات » باستعمال « تلوث الماء » و « البحيرات » لا « التلوث » و « البحيرات » .

وعلى مسئول البحث أن يكون متنبهاً دائماً لاحتمال الحشو في الاستراتيجية^(١) . وقد أوضحنا ذلك فعلاً على مستوى المصطلح في مثال بحوث الأسلحة ؛ فنظراً لأن المصطلح يعني « العسكرية » فإنه لا مبرر لربطه بأي مصطلح عسكري عام آخر (على الرغم من أننا قد نحتاج إلى مصطلح مخصص « كالأسطول » مثلاً ، إذا كان الأمر يتطلب اقتصار عملية البحث على فرع بعينه من العسكرية) . وهناك أيضاً ظاهرة « الحشو في مرصد البيانات » ؛ فالمصطلحات المستعملة في مرصد البيانات تأتي في سياق معين ، وفي مرصد بيانات مكرس أساساً لمجال التعليم ، فإنه قد يكون من المفترض أن يكون لمصطلح « مكفوف » علاقة إما بتعليم المكفوفين وإما بتعليم الناس كيفية مساعدة المكفوفين . أما في مرصد بيانات يتناول علم المكتبات ، فإنه غالباً ما يكون من المحتمل أن يكون مصطلح « مكفوف » متصلاً بالخدمات المكتبية الخاصة بالمكفوفين .

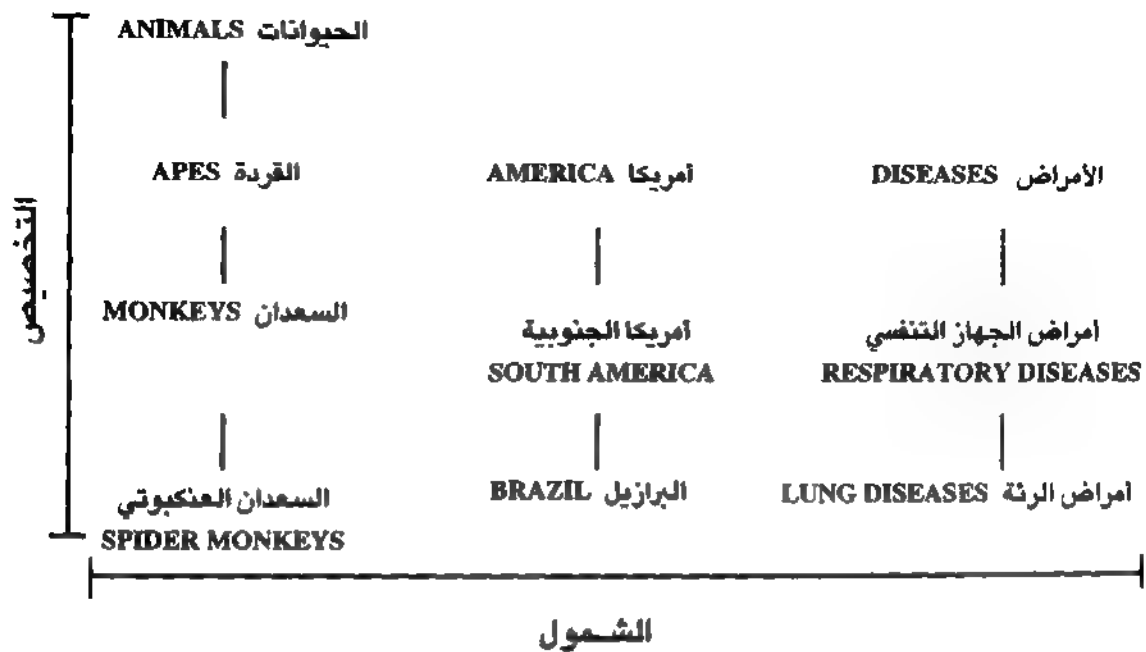
توسعة استراتيجيات البحث وتضييقها :

من الممكن التعبير عن نتائج البحث بناء على نسب الاستدعاء والتحقيق (انظر الفصل الثالث) . ويمكن لمسئول البحث ، في نطاق قيود معينة يفرضها النظام ، صياغة

(١) لقد نبه هارتر (1988) Harter أيضاً إلى ذلك .

استراتيجية تكفل الاستدعاء المرتفع أو التحقيق المرتفع أو نوعاً من المواءمة بين الاثنين. ومما لا شك فيه ، أنه من الممكن للربط بين مصطلحين بعلاقة « أو » أن يؤدي إلى استرجاع عدد من الوثائق يفوق ما يمكن استرجاعه بأي من المصطلحين منفرداً ، وأنه من الممكن للربط إذا كان له ما يبرره فعلاً ، أن يؤدي إلى الارتفاع بمستوى الاستدعاء. أما الربط بين المصطلحات اعتماداً على علاقة « و » أو علاقة « فيما عدا » فيمكن أن يؤدي إلى استرجاع عدد قليل من الوثائق، وبذلك يمكن أن يؤدي إلى ارتفاع نسبة التحقيق . وعادة ما تؤدي توسعة استراتيجية البحث بهدف الارتفاع بمستوى الاستدعاء إلى انخفاض نسبة التحقيق ، أما تضيق البحث بهدف الارتفاع بمستوى التحقيق فعادة ما يؤدي إلى خفض الاستدعاء .

وهناك العديد من النماذج التي توضح كيفية توسعة مجال البحث أو تضيقه . ومن بين السبل النظر إلى استراتيجية البحث باعتبارها ثنائية البعد ، والبعدان هنا هما الشمول والتخصيص . فإذا نظرنا إلى طلب للمعلومات حول « أمراض الرئة في السعدان العنكبوتي في البرازيل » والذي يمكن التعبير عنه على النحو التالي :



والشمول هو مدى الحاجة إلى جميع الأوجه الصريحة في طلب المعلومات ، في استراتيجية البحث (ويسمى ذلك أيضاً بمستوى الربط) ، أما التخصيص فهو مستوى التفصيل في التعبير عن وجه معين . ومدى التخصيص في استراتيجية البحث

يحكمه مدى التخصيص في اللغة المستعملة في مرصد البيانات . فما لم يكن مصطلح «السعدان العنكبوتي» موجوداً في مرصد البيانات ، فإنه لا يمكن البحث عن هذا الجانب بدقة ، وإنما يتعين الاكتفاء بما هو أكثر شمولاً ، وربما كان مصطلح السعدان . ومن الممكن التعبير عن الطلب بشكل مخصص وشامل في الوقت نفسه بالصيغة التالية :

السعدان العنكبوتي و البرازيل و أمراض الرئة .

وقد لا تسترجع هذه الاستراتيجية المحكمة وثائق كثيرة ، إلا أن كل ما تسترجعه يمكن أن يكون وثيق الصلة بموضوع الطلب . ومن الممكن توسعة البحث لاسترجاع المزيد من الوثائق ، على أمل الارتفاع بنسبة الاستدعاء ، وذلك بالحد من الشمول أو من التخصيص أو من كليهما معاً .

ومن الممكن على سبيل المثال ، القول جديلاً ، بأن أي وثيقة تتناول أمراض الرئة في السعدان البرازيلي يمكن أن تكون بوجه عام متصلة بشكل ما بموضوع الطلب ، ومن ثم فإنه يمكن توسعة البحث ليصبح :

السعدان و البرازيل و أمراض الرئة .

وينبغي أن نأخذ في حسابنا أن « التحرك إلى أعلى » في التفريع الهرمي للمصطلحات على هذا النحو (أي الانتقال من المصطلحات المخصصة جداً إلى الأقل تخصيصاً ، ينطوي على ما هو أكثر من إحلال مصطلح محل آخر ؛ فهو ينطوي على بحث عن الجنس كله generic search ، أي إحلال جميع المصطلحات اللازمة للتعبير عن الجنس بأكمله محل المصطلح المخصص . ويتطلب الأمر في هذه الحالة استعمال مصطلح «السعدان» وجميع المصطلحات الواردة في مرصد البيانات والتي تدل على أنواع معينة من السعدان (بما في ذلك السعدان العنكبوتي) . (ومن الممكن بالطبع الحد من مدى التخصيص في أي وجه ، حيث يمكن لمسئول البحث ، على سبيل المثال ، أن يفضل الارتفاع في التفريع الشجري من البرازيل إلى أمريكا الجنوبية ، بدلاً من الصعود من السعدان العنكبوتي إلى السعدان ، أو يتجاوز عن أحد مستويات هذا التفريع) . وإذا افترضنا أن مرصد البيانات يعتمد على لغة محكمة البنيان ، كما كنز مثلاً ، فإنه ينبغي أن تكفل برامج البحث المتسمة بالكفاءة إمكانية اشتغال الاستراتيجية على جميع

المصطلحات اللازمة لإجراء بحث شامل من هذا النوع (أي مصطلح « السعدان » وجميع المصطلحات المتفرعة عنه في الشجرة) بأمر واحد . وتسمى هذه الإمكانية في المدلاين بالقدرة على الإحاطة بكل العناصر explode .

أما الأسلوب الآخر لتوسعة البحث فهو الحد من الشمول في الاستراتيجية ؛ أي إسقاط أحد الأوجه . فمن الممكن القول جديلاً ، على سبيل المثال ، أنه يمكن لجميع المقالات التي تتناول أمراض الرئة في السعدان العنكبوتي ، سواء في البرازيل أو في غيرها ، أن تكون مفيدة لمن طلب المعلومات ، ومن ثم فإنه يمكن إسقاط الجانب الجغرافي كلية . ومن المفترض أن يؤدي ذلك إلى استرجاع عدد كبير من الوثائق ، وربما يؤدي إلى الارتفاع بنسبة الاستدعاء . ومما لا شك فيه أنه من الممكن الحد من التخصيص والشمول أو الإحاطة في استراتيجية البحث في الوقت نفسه .

ولننظر في الاستراتيجيات التالية وما لها من تأثير على الاسترجاع :

| عدد الوثائق المسترجعة | | الاستراتيجية |
|-----------------------|-----------------|--|
| غير المتصل بالموضوع | المتصل بالموضوع | |
| ٠ | ٢ | السعدان العنكبوتي و البرازيل و أمراض الرئة |
| ١ | ٥ | السعدان و البرازيل و أمراض الرئة |
| ١٢ | ٨ | السعدان و أمراض الرئة |

ومن الواضح هنا أثر الحد من كل من التخصيص والشمول على كل من الاستدعاء والتحقيق ؛ فقد بلغت أدق الاستراتيجيات الحد الأقصى للتحقيق ، إلا أنها لم تسترجع سوى ربع الوثائق المتصلة بالموضوع (٢ من ٨) ، بينما استرجعت أعم الاستراتيجيات المزيد من الوثائق المتصلة بالموضوع . (وربما كانت كل ما في مرصد البيانات) مع انخفاض في التحقيق إلى ٤٠ بالمئة فقط (٨ من ٢٠) .

ويحدد تصنيف شائع ، يصفه ماركي وكوشرين (Markey and Cochrane (1981 وغيرهما ، معالم أربعة أساليب رئيسة لبناء استراتيجيات البحث :

١ - أسلوب قوالب البناء .

٢ - أسلوب الأجزاء المتتابعة .

٣ - أسلوب تقديم أحد الأوجه المخصصة على ما عداه .

٤ - أسلوب « زراعة اللؤلؤ بالاستشهاد المرجعي » .

وأسلوب قوالب البناء هو المستعمل في أنحاء هذا الفصل ؛ ويقوم على التحقق من الأوجه التي تشكل مكونات البحث ، ثم التحقق من المصطلحات اللازمة للتعبير عن كل وجه تعبيراً كاملاً ، وأخيراً الربط بين هذه المصطلحات بالعلاقات المنطقية الملائمة كما في هذا المثال :

١ - البرازيل

٢ - السعدان العنكبوتي

٣ - أمراض الرئة

٤ - ١ و ٢ و ٣

وفي أسلوب الأجزاء المتتابعة ، يتم أولاً استرجاع مجموعة تمثل أحد أوجه البحث ، ثم العمل على تضيق حدود هذه المجموعة بإدخال الأوجه الأخرى الضرورية ، التي أمكن التحقق منها في مرحلة التحليل الموضوعي ، كما في هذا المثال :

١ - البرازيل

٢ - ١ و السعدان العنكبوتي

٣ - ٢ و أمراض الرئة

وهذا مجرد اختلاف طفيف عن أسلوب قوالب البناء ، ونتائج كلا الأسلوبين واحدة، ولا اختلاف إلا في تسلسل الاستراتيجية .

أما أسلوب تقديم أحد الأوجه المخصصة على ما عداه فأوقع من كل من أسلوب قوالب البناء وأسلوب الأجزاء المتتابعة . ويميز ماركي وكوشرين **Markey and Cochrane** (1981) بين هذا الأسلوب والأسلوب الذي يسميانه « تقديم أقل الأوجه في عدد التوقعات على ما عداه » (أي البدء بالوجه الذي يحتمل أن يسترجع أقل عدد من الوثائق) . وعلى الرغم من احتمال وجود بعض أمثلة لعمليات البحث التي لا يكون فيها أكثر الأوجه تخصيصاً هو في الوقت نفسه الوجه الذي يمكن أن يسترجع أقل عدد من الوثائق ، فإن

هذا الموقف لا يمكن أن يحدث في التطبيق الفعلي ، ولهذا فإنه لا مبرر للتمييز (*) . وبالنسبة لكل المقاصد والأهداف ، فإن أسلوب « تقديم الوجه المخصص على ما عداه » وأسلوب « تقديم أقل الأوجه في عدد التوقيعات » مطابقان في الأساس للأسلوب المسمى في الطبعتين السابقتين من هذا الكتاب ، بأسلوب « أقل العناصر شيوعاً » . وبالنسبة لكثير من عمليات البحث يمكن أن يكون هناك وجه يسترجع عدداً من الوثائق أقل بكثير مما تسترجعه الأوجه الأخرى . ويرتبط هذا الوجه أياً كان بحدود تغطية مرصد البيانات . ويبدو من المنطقي ، في المثال المستخدم هنا ، أن نتوقع أن يكون من الممكن للوجه الخاص بالسعدان العنكبوتي ، في كثير من مراصد البيانات ، أن يسترجع عدداً من الوثائق أقل مما يمكن أن يسترجعه الوجهان الآخران (البرازيل و أمراض الرئة) . ومن ثم فإن الأسلوب الملائم لإجراء عملية البحث هذه قد يكون البدء أولاً بالسعدان العنكبوتي . وربما كان من الممكن لهذا الأسلوب أن يسترجع عدداً قليلاً من الوثائق ، التي قد يرى المرء الاطلاع عليها جميعاً ، دون تقييد عملية البحث على أي نحو آخر .

ويختلف الأسلوب الذي يسميه ماركي وكوشرين (Markey and Cochrane 1981) بأسلوب (زراعة اللؤلؤ بالاستشهاد المرجعي) تمام الاختلاف عن الأساليب الأخرى ، وهو عبارة عن طريقة تكرارية يستخدم فيها مسئول البحث أيسر أو أقرب السبل المحتملة للعثور على وثيقة مفيدة أو أكثر ، ثم يتم الاطلاع على التسجيلات الوراقية لهذه الوثائق ، للتعرف على ما يرتبط بها من مصطلحات (مستقاة من لغة مقيدة أو كلمات واردة في العناوين أو المستخلصات) ثم يتم بعد ذلك إدخال المصطلحات الملائمة في الاستراتيجية. ^(١) ويمكن للاستراتيجية التي روجعت أن تسترجع حينئذ وثائق توجي بالمزيد من المصطلحات ، وهكذا إلى أن يصبح مسئول البحث راضياً عن النتائج . وهذا الأسلوب متمم فعلاً لغيره ؛ فمن الممكن على سبيل المثال اتباع أسلوب أقل التوقيعات للعثور على الوثائق القليلة الأولى ، كما يمكن ، بدلاً من ذلك بالطبع استرجاع الوثيقة أو الوثائق الأولى بخصائص « الوثيقة المعروفة » ، نظراً لأنها وثائق تتصل بالموضوع ، معروفة لطالب المعلومات من قبل .

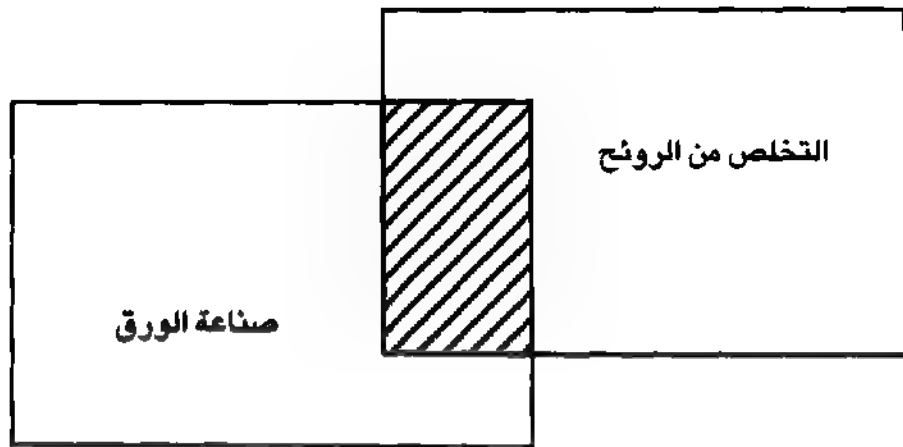
(*) يقصد التمييز بين أكثر الأوجه تخصيصاً والوجه الذي يمكن أن يسترجع أقل عدد من الوثائق . (المترجم) .

(١) ألقى الضوء على أهمية هذا الأسلوب مؤخراً لانكستر وزملاؤه (Lancaster et al. 1993) عند المقارنة بين طرق البحث المستخدمة من جانب المستفيدين من المكتبات وتلك التي يستخدمها المكتبيون المتمرسون في البحث في مراصد البيانات .

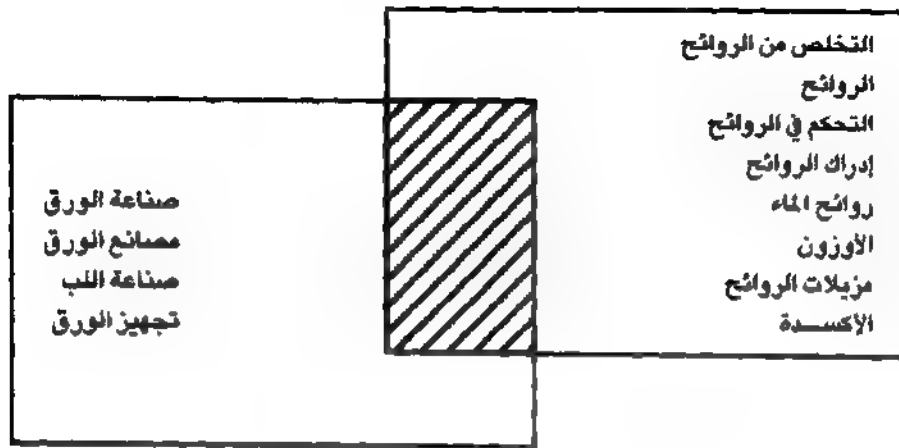
وفي تناوله لعملية البحث يميز بيتس (1979 b, 1987) Bates بين استراتيجية البحث وتكتيك البحث : فالأولى تتصل بالخطة العامة لإجراء عملية البحث بينما تتصل الثانية بالخطوات المحددة التي يتم اتخاذها عند إجراء البحث (كاختيار المصطلحات المزمع استعمالها على سبيل المثال) وتوسعة مجال البحث أو تضيقه ، والاستمرار في إجراءات البحث بوجه عام) . كذلك يحدد بيتس (1979 a) Bates معالم ما يسمى «التكتيكات النظرية» ، وهي في الواقع أساليب للتحليل الموضوعي لطلبات البحث عن الوثائق.

بعض عينات عمليات البحث :

من الأوفق دائماً تحليل طلب الحصول على المعلومات إلى أوجهه الأولية قبل الشروع في إجراء البحث . ويتبين من تحليل طلب الحصول على المعلومات حول التخلص من الروائح في صناعة الورق ، أنه يتكون من جانبين ، هما التخلص من الروائح وصناعة الورق ، ويتم البحث عن الوثائق بناء على الناتج المنطقي لهذين الجانبين :



وينبغي ترجمة ناتج التحليل الموضوعي إلى المصطلحات المستعملة للتعبير عن الموضوعات في مرصد البيانات المزمع إجراء البحث فيه . ومن الممكن التعبير عن مجموعات المصطلحات بيانياً على النحو التالي :



ويثير هذا المثال بعض النقاط المهمة ؛ فربما يكون مسئول البحث قد استهل التعامل مع لغة مرصد البيانات تحت مصطلح « التخلص من الروائح » ، ولا وجود لهذا المصطلح في لغة المرصد ، إلا أن الترتيب الهجائي يقود نحو « الروائح » و « التحكم في الروائح » و « إدراك الروائح » (*) ، وكلها يمكن أن تفيد كمصطلحات للبحث . واعتماداً على الإحالات الواردة بالمكنز يحال مسئول البحث من « الروائح » إلى « روائح الماء » ومن « التحكم في الروائح » إلى مزيلات الروائح . ويقود مصطلح « مزيلات الروائح » بدوره إلى « الأكسدة » و « الأوزون » اللذين يمكن أن يكونا مرتبطين بعملية إزالة الروائح . ويؤدي توسع مناظر للجانب الخاص بصناعة الورق بمسئول البحث إلى « مصانع الورق » و « صناعة اللب » و « تجهيز الورق » ، التي تبدو المصطلحات الوحيدة في مرصد البيانات هذا ، المتصلة بصناعة الورق . وإذا كان نظام الإحالات في المكنز مكتملاً فإن هذه المجموعة من المصطلحات تكون مكتملة أيضاً . أما إذا كان بناء المكنز غير مكتمل فإنه من الممكن ولا شك اقتراح مصطلحات بحث إضافية (اعتماداً على التسجيلات المسترجعة) أثناء إجراء عملية البحث . وعلى الأساس نفسه يمكن للقصور في بناء المكنز أن يؤدي بمسئول البحث إلى استعمال بعض المصطلحات غير الملائمة؛ فقد لا يكون كل من « الأكسدة » و « الأوزون » على سبيل المثال من الاختيارات المناسبة لهذا الموضوع بالذات .

ومن الممكن النظر إلى الشكل البياني الذي يشتمل على المصطلحات الكشفية باعتباره يعبر عن استراتيجية بحث معقولة مناسبة للطلب ؛ فإذا كانت هناك وثيقة قد

(*) ترد كل هذه المصطلحات ، والآخران منها مضاف ومضاف إليه تحت كلمة روائح ODORS تبعاً لقواعد النظم الإنجليزية . (المترجم) .

كشفت تحت واحد على الأقل من مصطلحات صناعة الورق وواحد على الأقل من مصطلحات التخلص من الروائح، فإنها يمكن أن تكون متصلة ، على نحو ما ، بالتخلص من الروائح في صناعة الورق .

ولا تتمتع جميع المصطلحات بالطبع بالوزن نفسه في علاقتها بطلب المعلومات ؛ فعلى الرغم من أن المصطلحات الواردة في الجانب الأيسر للشكل البياني متساوية جميعها في صلاحيتها للدلالة على صناعة الورق ، فإن تلك المصطلحات الواردة في الجانب الأيمن ليست على قدم المساواة من حيث اتصالها بالتخلص من الروائح. ويبدو مصطلح « التحكم في الروائح » هو الأوثق صلة بالموضوع ، وربما يصاحبه المصطلحات التي يمكن أن تعبر عن بعض أساليب معينة للتحكم في الروائح، كالاكسدة، ومزيلات الروائح ، والأوزون ، في المستوى التالي . وتقع مصطلحات «الروائح» و«روائح الماء» و« إدراك الروائح » في مرتبة أدنى في الترتيب الطبقي نظراً لأنها لا تعني مباشرة عنصر التخلص من الروائح . إلا أنه من الممكن لوثيقة مكشوفة تحت واحد من هذه المصطلحات ، وتحت أحد مصطلحات صناعة الورق أيضاً ، أن نتوقع لها ، وبشكل معقول ، الاتصال إلى حد ما بموضوع التخلص من الروائح في صناعة الورق . ولفكرة إمكانية ترتيب قائمة المصطلحات طبقياً وفقاً لاحتمال اتصال المصطلحات بالموضوع أهميتها، نظراً لأن هذا التتابع أو التسلسل يعبر عن التسلسل الذي يمكن به منطقياً إجراء عملية البحث :

١- أ

٢- أ و ب

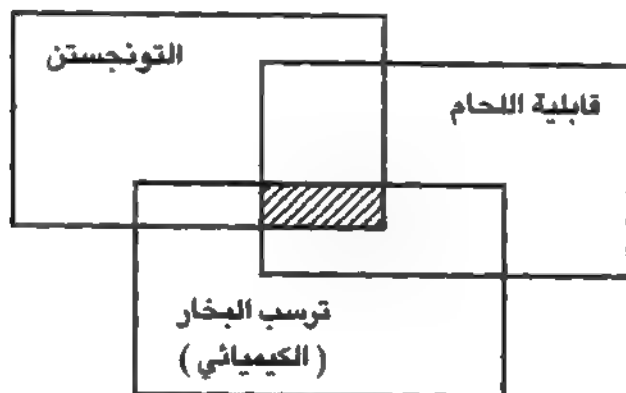
٣- أ و ج

٤- أ و د

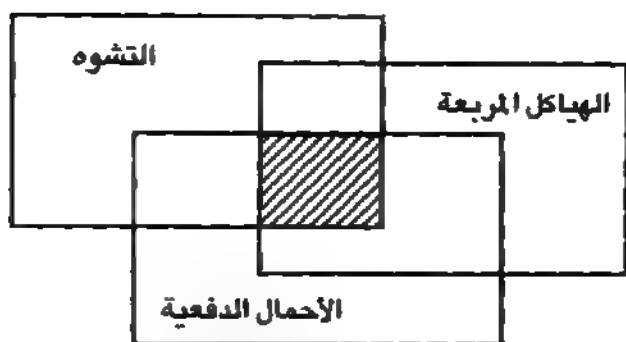
حيث يعبر « أ » عن مصطلحات صناعة الورق ، بينما يعبر « ب » و « ج » و « د » عن مصطلحات الروائح ، في تتابع يتفق واحتمالات الاتصال بالموضوع . وللمستفيد الذي يريد العثور على أقرب الوثائق صلة بالموضوع فقط ، فإنه يمكن إنهاء البحث عند الصيغة الثانية للبحث ، على افتراض أنه قد أمكن العثور على بعض الوثائق المتصلة بالموضوع . أما المستفيد الذي يحتاج إلى وراقية شاملة ، فإنه قد يرغب في مواصلة البحث، لكي يطمئن إلى أنه لم يفته شيء .

هذا ، ومن الممكن لأي حاجة إلى المعلومات ، مهما بدت معقدة ، أن تترجم إلى استراتيجية بحث منطقية ، طالما كان من الممكن إيجازها بالطريقة التي أوضحناها . وفيما يلي بعض الأمثلة على المستوى المفاهيمي أو النظري :

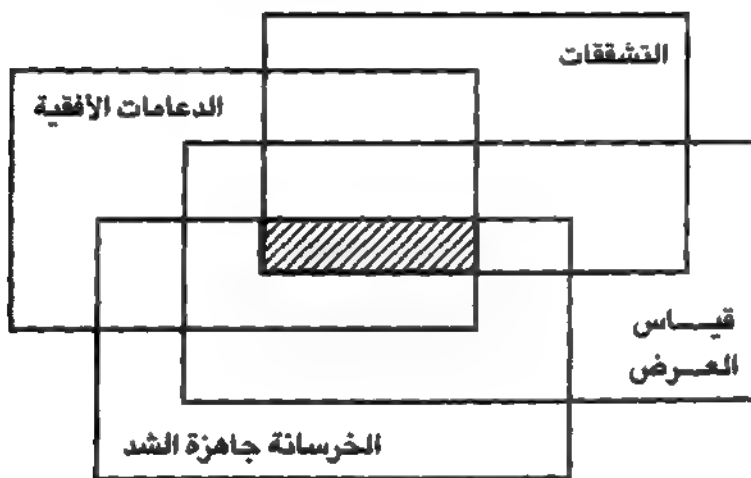
قابلية التونجستن (*) المترسب عن البخار الكيميائي للحام :



تشوه الهياكل المربعة تحت الأحمال الدفعية الموزعة :

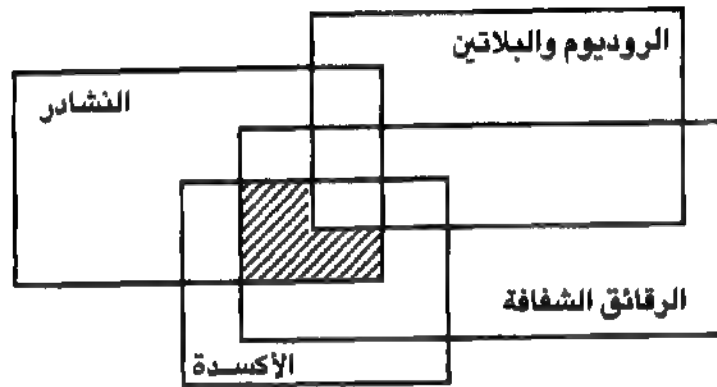


قياس عرض التشققات في الدعامات الأفقية المسلحة جاهزة الشد جزئياً :



(*) عنصر فلزي يستخدم في تقوية الفولاذ . (المترجم) .

استخدام الرقائق الشفافة فيما عدا رقائق الروديوم والبلاتين في أكسدة النشادر :



ويتعين الآن التوسع في كل طريقة من هذه الطرق الخاصة بالتعبير عن الحاجة إلى المعلومات ، في لغة مرصد البيانات المزمع البحث فيه . وكما سبق أن أشرنا ، فإنه قد لا يكون من الممكن دائماً التعبير عن موضوع ما باللغة المحددة لمرصد البيانات ؛ فقد لا يكون من الممكن على سبيل المثال التعبير تحديداً عن ترسب البخار « الكيميائي » ، أو تمييز الهياكل المربعة عن غيرها من الأنواع ، أو التعبير عن « قياس » العرض . وواقع الأمر أنه من الممكن لأي من أوجه الحاجة إلى المعلومات أن يكون غائباً تماماً في لغة مرصد البيانات ، مما يضطر المستفيد لإجراء بحث أعرض من حاجته . وما لم يكن هناك مصطلح يعبر عن الاحتمال فإنه من الممكن للمرء البحث عن تشوه الهياكل المربعة فقط . وإذا كان هناك مصطلح مخصص للأحمال ، وليس هناك مصطلح للتشوه ، فإنه من الممكن البحث عن الاحمال الدفعية على الهياكل المربعة فقط ، وهكذا .

هذا ، وقد سبق أن أشرنا أيضاً إلى أنه قد لا تكون هناك علاقة تطابق مباشر بين الموضوعات التي أمكن التحقق منها ولغة مرصد البيانات ؛ فمن الممكن أحياناً للمصطلح الواحد في اللغة أن يعبر عن وجهين أو أكثر من أوجه البحث . ومن الأمثلة على ذلك مصطلح «أكسدة النشادر» أو مصطلح « الدعامات الأفقية المسلحة جاهزة الشد » . أضف إلى ذلك أنه من الممكن للغة مرصد البيانات أن تقسم المجالات الموضوعية بطريقة تختلف عن التحليل الموضوعي الذي يقوم به مسئول البحث . كذلك يمكن ، وإن كان ذلك من المستبعد ، أن تشتمل اللغة على مصطلح « عرض التشققات » وكذلك على مصطلح « القياس » . وعلى المرء أن يكون حريصاً في التعرف على هذه الاحتمالات عند ترجمة استراتيجية البحث من مستوى الموضوعات إلى مستوى المصطلحات .

وليس من الضروري تصوير استراتيجية البحث بيانياً ، تماماً كما أوضحنا ، وإن كان من المفضل ولا شك اتباع أسلوب منطقي إلى حد ما في تحليل الحاجة إلى المعلومات ، قبل إجراء عملية البحث . ومما لا شك فيه ، أن استعمال الرسومات البيانية من النوعية التي بينها هنا ، والتي يمكن رسمها تخطيطياً في بضع ثوانٍ ، يمكن أن يساعد في الدلالة على المطلوب في البحث . وقد أكد الأسلوب جدواه في الاستخدام من جانب اختصاصي المعلومات ، عند مناقشة الحاجة إلى المعلومات مع المستفيد (انظر على سبيل المثال . (Smith, 1976

ويتمتع المرء بقدر كبير من المرونة في صياغة استراتيجية للبحث في مرصد البيانات على الخط المباشر ؛ فإذا كان المستفيد بحاجة إلى عدد قليل فقط من الوثائق المتصلة بالموضوع فإنه يتعين اتباع أكثر الأساليب تحديداً ووضوحاً ، أما إذا كان بحاجة إلى بحث شامل فعلاً ، فإنه يتعين توسعة مدى الاستراتيجية لتجنب ضياع وثائق قيمة . ويمكن للتخصيص المفرط في البحث أن يكون خطراً فعلاً . وأمامنا البحث الخاص بقياس عرض التشققات ؛ فإذا تم الربط بين المصطلحات الأربعة ، « الدعامات الأفقية » و « التشققات » والخرسانة جاهزة الشد » و « القياس » بعلاقة « و » المنطقية ، فإنه من المؤكد تقريباً أن تكون جميع الوثائق المسترجعة متصلة بالموضوع . إلا أنه يمثل هذه الاستراتيجية المخصصة فإنه من المحتمل جداً ألا يتم استرجاع كل ما يمكن أن يكون متصلاً بالموضوع . بل إننا قد لا نسترجع فعلاً شيئاً على الإطلاق . فمستول البحث يفترض أن جميع الوثائق المتصلة بالموضوع قد كشفت بدقة على النحو الذي يتصوره ، وهذا افتراض محفوف بالمخاطر .

ويمكن للاستراتيجية المكونة « من الدعامات الأفقية » و « التشققات » و « الخرسانة جاهزة الشد » ، والتي تستبعد مصطلح « القياس » أن تسفر عن شيء يتصل بالموضوع ؛ فربما تسترجع مقالات عن قياس التشققات لم يستعمل مصطلح « القياس » في تكثيفها ، كما يمكن أن تسترجع أيضاً تقارير عامة عن تشقق الدعامات الأفقية الخرسانية جاهزة الشد ، والتي تتناول قياس التشققات حتى وإن لم يحظ هذا الجانب بالتغطية على وجه التحديد في التكثيف . وواقع الأمر أنه من الممكن لأي من التجميعات المؤتلفة للمصطلحات والموضحة في الشكل رقم (٢٧) أن تسفر وبشكل معقول عن بعض الإشارات الوراقية المتصلة بموضوع الاهتمام المخصص .

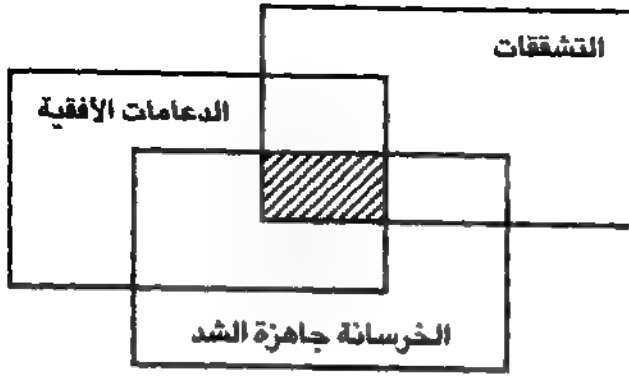
ومن اليسير جداً اختبار مختلف التجميعات المؤتلفة للمصطلحات الواردة في الشكل رقم (٣٧) في البحث على الخط المباشر . ويمكن للمستفيد أن يجرب أكثر السبل تحديداً في البداية، وما لم يكن هذا الأسلوب ناجحاً ، فإنه يوسع البحث على نحو منطقي . وحتى في حالة ما إذا كان البحث المحدد ناجحاً ، أي في حالة ما إذا استرجع بعض الوثائق المتصلة بالموضوع ، فإنه يمكن لمسئول البحث أن يعمم على المستويات الأخرى الممكنة ، إذا كان المستفيد بحاجة إلى نتائج شاملة فعلاً .

وهناك بعض السلبيات أو المحاذير المرتبطة بالتعميم أيضاً ؛ فكلما كان البحث عاماً ازدادت احتمالات استرجاع المزيد من الوثائق غير المتصلة بالموضوع . فلا يمكن للاستراتيجية عالية التحديد المكونة من « الدعامات الأفقية » و « الخرسانة جاهزة الشد » و « القياس » و « التشققات » أن تسترجع وثائق لا تتصل بالحاجة الراهنة إلى المعلومات ، إلا أنها قد لا تسترجع أية وثائق على الإطلاق . وكلما عممنا بدءاً من هذه النقطة ازداد عدد الوثائق التي يحتمل استرجاعها ، فضلاً عن ازدياد احتمال العثور على بعض الوثائق المتصلة بالموضوع . إلا أنه تزداد في الوقت نفسه احتمالات استرجاع وثائق لا تتصل بالموضوع . ويمكن للاستراتيجية المكونة من « الخرسانة جاهزة الشد » و « التشققات » أن تسترجع الكثير من الوثائق التي لا علاقة لها على وجه التحديد بقياس التشققات ، والأقل اتصالاً بقياس تشققات الدعامات الأفقية . إلا أنه يمكن في الوقت نفسه لهذه الاستراتيجية العامة أن تكون هي الاستراتيجية الوحيدة التي يمكن أن تسترجع الوثائق المتصلة على أي نحو كان ، بموضوع الاهتمام .

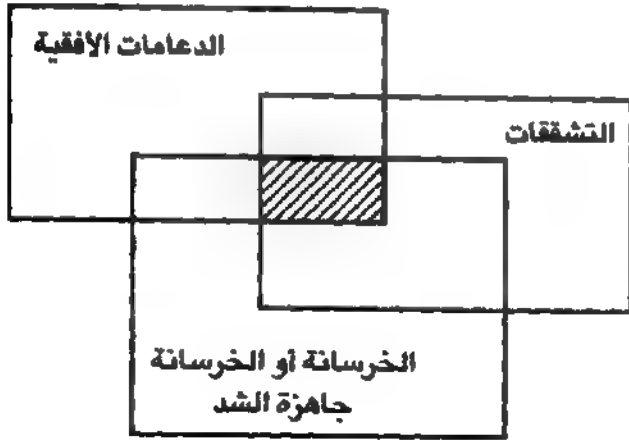
نخلص من كل ما سبق إلى أنه يتعين على المستفيدين أن يسلكوا أيسر السبل عند البحث في نظم استرجاع المعلومات . ومن يمن الطالع أن نظم الخط المباشر تيسر ذلك إلى حد بعيد . ولهذا السبب فإن البحث على الخط المباشر يمكن أن يكون هو الأكثر فعالية من حيث العثور على الوثائق المتصلة بالموضوع ، والأكثر كفاءة من حيث الاقتصاد في الوقت، وذلك إذا ما سار وفقاً للتسلسل التالي :

١ - ترجمة ناتج التحليل الموضوعي للحاجة إلى المعلومات إلى أوجهها المختلفة ، ذهنياً على الأقل ، وبيانياً أيضاً إذا كان ذلك مفيداً .

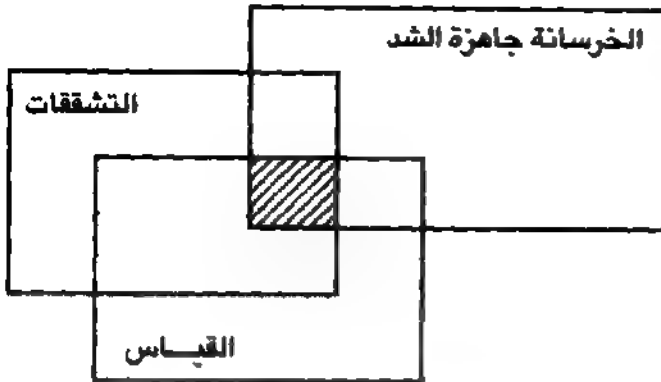
٢ - الاعتماد على قطاعات من لغة المرصد على الخط المباشر أو في شكلها المطبوع،



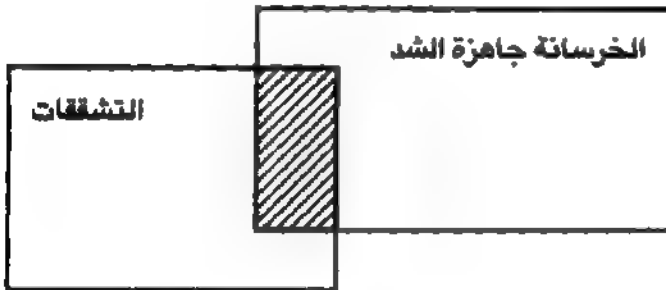
يمكن لموضوع القياس أن يكون قد حظي بالمعالجة فعلاً إلا أنه لم يكشف على وجه التحديد



يمكن لتقرير عن التشققات في الدعامات الأفقية الخرسانية بوجه عام ، أن يكون متصلاً بموضوع التشققات في الدعامات الأفقية الخرسانية جاهزة الشد ، كما يمكن لمقالة عن الخرسانة جاهزة الشد أن تكون قد كشفت خطأ تحت المصطلح الأعرض « خرسانة »



يمكن لتقرير عن قياس التشققات في الخرسانة جاهزة الشد أن يكون متصلاً وبالقدر نفسه بقياس التشققات في الدعامات الأفقية الخرسانية.



يمكن لأي تقرير عن تشقق الخرسانة جاهزة الشد أن يناقش قياس التشققات .

الشكل رقم (٣٧) : أساليب البحث المحتملة لموضوع قياس عرض التشققات في الدعامات الأفقية الخرسانية سابقة الشد جزئياً .

للتعرف على المصطلحات المتصلة بكل وجه في اللغة الخاصة بمرصد البيانات المزمع البحث فيه .

٣ - تجربة التجميعات المؤتلفة للمصطلحات في تسلسل منطقي ، تبعاً لاحتمال اتصالها بطلب الحصول على المعلومات .

٤ - عرض بعض الإشارات الوراقية المسترجعة لمعرفة ما إذا كانت هذه التسجيلات توحى بمصطلحات بحث أو أساليب بحث إضافية .

٥ - طباعة نتائج البحث عندما يصبح هذا الإجراء التكراري غير مجدٍ ، أي عندما يتوقف عن الإيحاء بمصطلحات بحث أو أساليب بحث جديدة .

وهناك بعض سبل ضغط جهد البحث أو وقت البحث إلى أدنى حد ممكن ، في كثير من نظم الخط المباشر . ويمكن لإمكانية البتر إذا ما استخدمت بحرص ، أن تقلل إلى حد بعيد من مقدار التعامل مع لوحة المفاتيح الذي كان من الممكن أن تدعو الحاجة إليه ما لم نلجأ إلى هذا الأسلوب . فإذا كان هناك على سبيل المثال قائمة مطولة بمصطلحات صناعة الورق في أحد مراصد البيانات ، فإنه قد يكون من الممكن إدخال معظم هذه المصطلحات في استراتيجية البحث بالبتر البسيط ورق : PAPER .

ومن بين ما يوصى به أيضاً ترتيب عناصر استراتيجية البحث بحيث يبدأ التعامل مع الجانب الذي نتوقع له أقل عدد من الوثائق في المرصد قبل غيره ؛ فإذا كان لدينا على سبيل المثال بحث عن المعلومات حول لحام أنواع الفولاذ الكروم والنيكل والفاناديوم ، فإنه من الممكن أن يكون هناك في لغة النظام عشرون مصطلحاً تعبر عن فكرة اللحام ، ومصطلح واحد فقط يدل على كل نوع من أنواع الفولاذ موضوع الاهتمام . ومن الأوقع بكل المقاييس البدء بمصطلح « فولاذ » أولاً ، نظراً لأنه قد لا يكون هناك بمرصد البيانات تحت هذا المصطلح سوى عدد من الوثائق لا يتجاوز أصابع اليدين . وإذا كان الأمر كذلك فعلاً فإنه قد يكون من الأبسط والأيسر والأسرع عرض كل هذه التسجيلات على الخط المباشر ، ثم ننتقي منها تلك التي تتناول اللحام . وأسوأ أسلوب للتعامل مع هذا الطلب هو البحث عن كل مصطلحات اللحام ، والتي يمكن أن تسترجع عدة مئات من التسجيلات ، ثم تعديل ذلك بنوع الفولاذ موضوع الاهتمام . وليس من الممكن دائماً التعرف على العنصر الأقل شيوعاً في البحث ، إلا أن ذلك يحدث في غالب الأحيان .

وربما يكون من الممكن تقييد البحث باستعمال الكلمات التي ترد في العناوين أو المستخلصات ، وذلك بعد تضيق الاستراتيجية بالبحث في المصطلحات المقيدة لمصدر البيانات . وتتيح هذه إمكانية إجراء بحث أكثر تخصيصاً مما يمكن أن تسمح به اللغة المقيدة الخاصة بمصدر البيانات . فمن الممكن على سبيل المثال أن تقود المصطلحات المقيدة مسئول البحث نحو مجموعة من الوثائق التي تتناول تشقق الخرسانة جاهزة الشد، إلا أن هذه المجموعة تضم عدداً كبيراً من الوثائق . وليس هناك مصطلح مقيد للدعائم الأفقية ولا للقياس ، إلا أنه من الممكن تضيق عملية البحث أكثر بالبحث في العناوين أو المستخلصات الخاصة بالوثائق التي تضمها المجموعة لمعرفة ما إذا كان مصطلح « الدعائم الأفقية » أو أي لفظ يدل على القياس موجوداً .

ويمكن لبعض مراصد البيانات أن تتمتع بحيل خاصة يمكن الاعتماد عليها للارتفاع بمستوى اتصال مخرجات البحث بالموضوع ، وذلك بتضيق البحث ، الذي كان من الممكن ، ما لم يضيق ، أن يسترجع عدداً كبيراً من الوثائق . ومن بين هذه الحيل بوجه خاص إمكانية وزن المصطلحات الكشفية التي تخصص للوثائق ، للتعبير عن أهميتها النسبية للموضوع الذي تتناوله . ويمكن للوزن أن يتم ببساطة بميزان من مستويين: مصطلحات أكثر أهمية ومصطلحات أقل أهمية . وتميز المصطلحات الموزونة في الملفات برمز خاص، قد يكون نجمة (*) . وبذلك يمكن لمصطلح « التلوث الحراري »* ألا يسترجع إلا تلك الوثائق التي تتناول موضوع التلوث الحراري تفصيلاً . وتجنب الوثائق التي تهتم بالموضوع اهتماماً عارضاً . ومن الممكن ربط المصطلح الموزون بغيره من المصطلحات لصياغة استراتيجية بحث على النحو المعتاد .

بحث النصوص الحرة :

تكفل كثير من مراصد البيانات التي يمكن التعامل معها الآن على الخط المباشر أو في شكل أسطوانات ضوئية مكتنزة ، للمستفيد إمكانية البحث في النصوص الحرة (أي اللغة الطبيعية ، الكلمات أو العبارات التي ترد في جميع الحقول النصية ، بما في ذلك العناوين والمستخلصات ، والنصوص الكاملة للوثائق) . وهناك بعض مراصد البيانات التي لا تتيح التعامل إلا عن طريق هذه النصوص ، بينما توفر مراصد أخرى إمكانية التعامل مع كل من النصوص الحرة والمصطلحات المقيدة .

ويستفيد بحث النصوص الحرة كما يستفيد غيره من أنواع البحث من الأسلوب المنطقي للتحليل الموضوعي . ومن أوجه الاختلاف الجوهرية أن ناتج التحليل الموضوعي لا يترجم إلى مجموعة بعينها من المصطلحات المقيدة ، حيث يتعين على مسئول البحث ، بدلاً من ذلك ، أن يقرر أي الكلمات أو العبارات يمكن أن تميز الوثائق التي يمكن أن تتصل بالموضوع مما عداها . ووجه آخر من أوجه الاختلاف الجوهرية أنه من الممكن للبحث في اللغة الطبيعية أن يعتمد على تجميعات مؤلفة مرنة من الكلمات المفردة والعبارات ، تحت السيطرة الكاملة لمسئول البحث . والكلمة هي الوحدة الأساسية للبحث . والبحث في اللغة الطبيعية أصعب في بعض جوانبه من البحث في مرادف البيانات المعتمدة على اللغات المقيدة ، إلا أنه من ناحية أخرى يكفل عدداً من المزايا .

وأبرز مزايا اللغة الطبيعية أنها تتيح إمكانية إجراء عمليات البحث على مستوى من التخصيص لا حدود له . فمن الممكن على سبيل المثال البحث عن الوثائق التي يرد بها ذكر أسماء شركات أو منتجات أو عمليات أو أشخاص بعينهم . واستعمال معادلات برجر Berger في تحليل الذبذبات ، ومشروع ناسا لانجلي NASA - Langley للطاقة الشمسية ، وآلات الطحن التي تنتجها شركة Schneider Maschinenbau GmbH ، وتصميم وبناء جسر بوسبورس Bosphorus Bridge ، كل هذه أمثلة مناسبة لاحتياجات المعلومات التي يمكن تلبيتها بسرعة فائقة بالبحث في النصوص الحرة . ومن الممكن إجراء عمليات البحث هذه في نظم اللغات المقيدة أيضاً ، إلا أن ذلك يمكن أن يكون أكثر صعوبة .

ولبدء التماس أقل العناصر شيوعاً في البحث أهميته الخاصة في استعمال اللغة الطبيعية ؛ فمن الممكن على سبيل المثال التعامل مع طلب المعلومات حول استعمال معادلات برجر في تحليل الذبذبات ، وبشكل فعال جداً ، وذلك بالبحث بساطة عن كلمة BERGER دون سواها . ومن الممكن لاستعمال الكلمة أن يضيق البحث ويقصره على عدد من التسجيلات لا يتجاوز أصابع اليدين ، يمكن عرض هذه التسجيلات لمعرفة ما إذا كان أي منها يتناول تطبيق معادلات برجر في تحليل الذبذبات . وليس من الضروري حتى التفكير في الأسماء المركبة في أشكالها الكاملة . كذلك يمكن على سبيل

المثال ، لكلمة الطحن GRINDING مرتبطة بكلمة شنايدر SCHNEIDER أن تكون ملائمة للتعامل مع طلب المعلومات المتعلقة بآلات الطحن التي تنتجها شركة شنايدر . Schneider Maschinenbau GMBH

وحتى عند البحث عما لا يدخل ضمن الأسماء، فإنه غالباً ما يكون من الممكن اختزال البحث في اللغة الطبيعية في عدد قليل من الكلمات « المفتاحية » (أقل العناصر شيوعاً). وفيما يلي بعض الأمثلة ، حيث نبين موضوع الحاجة إلى المعلومات ونسجل في مقابله الاختزال المحتمل للحاجة في بعض التجميعات المؤتلفة من الكلمات ، والتي يمكن أن تكون قادرة على التمييز بما فيه الكفاية لإجراء البحث على الخط المباشر :

١ - المحطات الخرسانية الطافية للاستخدام

في المنطقة القطبية الشمالية الطافية و المنطقة القطبية الشمالية

٢ - عملية الضغط المتوازن الساخن

المتوازن

٣ - التصنيع الآلي للمعادن بليزر

نيوديميوم نيوديميوم

٤ - أثر مقاومات التأكسد على

زيت النخيل المكرر النخيل و مقاوم الأكسدة

وهكذا يتضح أنه من الممكن اختزال البحث إلى أقصى حد باختيار كلمة مفتاحية واحدة أو اثنتين فقط . وعلى مسئول البحث أن يتطلع دائماً لأبسط السبل وأكثرها مباشرة لاسترجاع الإشارات الوراقية المتصلة بالموضوع . وعادة ما يعني ذلك اختيار الكلمة الأساسية التي يمكن أن تنطبق على أقل عدد من الوثائق في الملف . فكلما « ساخن » أو « الضغط » يمكن أن ينطبقا على عدد كبير من الوثائق ، أما « المتوازن » فقد لا تنطبق إلا على عدد قليل ، ولهذا فإنها تعد الاختيار الأرجح للسبيل الأول لاسترجاع الإشارات الوراقية المتصلة بموضوع الضغط المتوازن .

وليست جميع عمليات البحث يسيرة التداول كتلك المستخدمة في هذه الأمثلة . فكلما كان موضوع البحث أقل تخصيصاً كان من الصعب تداوله اعتماداً على اللغة

الطبيعية ، نظراً لصعوبة التحقق من جميع الكلمات التي يمكن أن تعبر عن بعض الأوجه العامة للطلب. ولأسلوب الذي سبق أن أوضحناه ، والخاص بتحليل الحاجة إلى المعلومات إلى الأوجه التي تتكون منها ، ثم اختيار أفضل المصطلحات للتعبير عن كل وجه ، أهميته في بحث اللغة الطبيعية ، كما هو الحال في بحث اللغات المقيدة ، ووجه الاختلاف الحقيقي الوحيد أنه لا يمكن في بحث اللغة الطبيعية أن يكون هناك مكنز بنظام للإحالات يساعد مسئول البحث في اختيار جميع المصطلحات الملائمة . ففي نظام اللغة الطبيعية يتوقف الكثير على براعة المسئول عن البحث .

لنفترض أن هناك من يبحث عن المعلومات حول مستويات الزئبق في الماء ، حيث يمكن الاقتصار بالنسبة للوجه الخاص بالزئبق على كلمتي « زئبق MERCURY » و« زئبقي MERCURIC » إلا أنه من غير الملائم على الإطلاق اقتصار الوجه الخاص بالماء على مصطلح « ماء » فقط ؛ فهناك طرق كثيرة يمكن بها التعبير عن « الماء » في الوثائق، كالماء ، والبحر ، والمحيط ، والجدول ، والبحيرة ، والنهر ، وهكذا . ولإجراء بحث شامل في هذا الموضوع فإنه يتعين تدبر جميع المصطلحات التي يمكن أن تعبر عن الجانب الخاص بالماء ، في مرصد البيانات المزمع إجراء البحث فيه . وهذه ليست بالمهمة المستحيلة ، إلا أنها يمكن أن تكون أصعب مما هي عليه في نظم اللغات المقيدة ، والتي يمكن أن نتوقع فيها جميع المصطلحات الخاصة بالماء مرتبطة ببعضها البعض بنوع من الإحالات .

ويمكن ، من ناحية أخرى ، للبحث في اللغة الطبيعية ، أن يكون أيسر ، على أكثر من نحو ، من البحث في اللغة المقيدة ؛ فنظراً لأن المرء غالباً ما يتعامل على مستوى الكلمة ، فإنه عادة ما يكون من الممكن اختزال البحث في عدد من العناصر ، أقل مما هي عليه في بحث اللغة المقيدة . فمن الممكن على سبيل المثال اختزال طلب المعلومات المتصلة بالتخلص من الروائح في صناعة الورق ، والذي اعتمدنا عليه في مثال سابق ، على النحو التالي :



ومن الجدير بالذكر أنه على الرغم من أن الكلمة في حد ذاتها قد تكون غامضة أو غير محددة ، فإنه من الممكن لهذا الغموض أو الافتقار إلى التحديد أن يختفي تقريباً عندما ترتبط الكلمة بغيرها من الكلمات . فكلمة *beam* في الإلكترونيات تعني شيئاً مختلفاً تماماً عن معنى *beam* في الهندسة المعمارية (*) . إلا أنه يمكن للمجموعة المؤتلفة المكونة من *BEAM and (CRACK OR CRACKING)* أن تسترجع الوثائق التي تتناول الدعامات المعمارية الأفقية ، لا تلك التي تتناول أشعة الإلكترونات . وكذلك الحال أيضاً بالنسبة لكلمة *grain* وكلمة *migration* ، حيث لكل منهما عدة سياقات محتملة ، إلا أنهما عندما يتم الربط بينهما بعلاقة « و » يمكن أن يسترجعا الوثائق المتصلة بحركة الذرات في المعادن .

ومن قبيل المبالغة في التبسيط الحديث عن الكلمة باعتبارها الوحدة الأساسية في البحث في مراصد بيانات اللغة الطبيعية . وربما كان من الأفضل النظر في استراتيجية البحث بناء على كل من الكلمات وأجزاء الكلمات . فمعظم نظم الخط المباشر ونظم الأسطوانات الضوئية المكتنزة تكفل للمستفيد البحث بأجزاء الكلمات باستخدام البتر ، الذي يمكن أن يتخذ أربعة أشكال :

- ١ - البتر الأيمن ، أي إسقاط نهاية الكلمة؛ فمن الممكن للبتر: *CRACK* على سبيل المثال أن يسترجع الوثائق التي تشتمل على *CRACKED* و *CRACK* و *CRACKS* و *CRACKING* .
- ٢ - البتر الأيسر ، أي إسقاط بداية الكلمة ؛ فمن الممكن على سبيل المثال للبتر *MYCIN* : استرجاع الوثائق التي تتناول مجموعة كبيرة من المضادات الحيوية .
- ٣ - بتر الجانبين الأيسر والأيمن معاً .
- ٤ - بتر الحشو، أي تحديد بداية الكلمة ونهايتها وترك وسطها دون تحديد (مثل: *TRI ... COBALATE*) .

والبتر الأيمن هو الأجدى في كثير من التطبيقات ، وهو أداة قوية للبحث وفقاً لمجموعات كبيرة من الكلمات المتصلة ببعضها البعض . والبتر وسيلة مفيدة للاقتصاد

(*) في الأول تعني شعاع وفي الثانية تعني دعامة أفقية . (المترجم) .

في الوقت ، نظراً لأنه يتجنب الحاجة إلى إدخال قائمة من المصطلحات المشتركة في الجذر نفسه، كل على حدة . إلا أننا ينبغي ألا ننسى أنه من الممكن للبتر أن يفضي إلى بعض الكلمات التي لا علاقة لها بطلب معين على الإطلاق؛ فمن الممكن لـ : DOG أن يؤدي إلى استرجاع وثائق عن doggerel (أي الشعر الهزلي) أو مصرف Dogwood Bank أو عن الكلاب أيضاً (على الرغم من أن كل هذه الموضوعات قد لا ترد معاً في مرصد البيانات نفسه). والبتر : CAT غامض أيضاً (حيث يمكن أن يكون CATS, CAT, CAT- CRACK . إلا أنه من الممكن لكل من CAT و : CRACK أن يكونا دقيقين تماماً في استرجاع الوثائق المتصلة بالتشقق بالحفز Catalytic Cracking (*) .

البحث بالمصطلحات الموزونة :

كانت عمليات البحث في مرصد البيانات الإلكترونية قبل انتشار استخدام نظم الخط المباشر ، تتم بالتدقيق تسلسلياً في التسجيلات الوراقية المختزنة على أشرطة ممغنطة، حيث كان يتم إجراء عدة عمليات بحث في الوقت نفسه (التجهيز على دفعات) . وكان كثير من مراكز المعلومات يستخدم أسلوب « المصطلح الموزون » بدلاً من الأسلوب البوليني المؤلف في البحث ، وخاصة لعمليات البحث التي تتم لأغراض البث الانتقائي للمعلومات .

ولا يختلف منطق البحث بالمصطلحات الموزونة على منطق الجبر البوليني ؛ فوزن المصطلحات يتم لمحاكاة الاستراتيجية البولينية ، كما كان من الممكن استخدام كلا الأسلوبين تبادلياً ؛ فعملية البحث التي يتم إجراؤها بالمصطلحات الموزونة يمكن أن تتم أيضاً بالجبر البوليني ، وعملية البحث التي يتم إجراؤها بالأسلوب الثاني يمكن إجراؤها أيضاً بالأسلوب الأول . وفي البحث بالمصطلحات الموزونة كما في الأنواع الأخرى من أساليب البحث ، فإن الخطوة الأولى هي التحليل الموضوعي للطلب إلى الأوجه التي يتكون منها ، ثم التوسع في كل وجه باختيار المصطلحات الملائمة من لغة مرصد

(*) للبتر في العربية قصة تختلف عن قصته في الإنجليزية ؛ فالأخيرة غروية أو لصقية في نظامها الصرفي ، شأنها في ذلك شأن بقية اللغات الهند أوروبية ، أما الأولى فعالية الاشتقاق . وإذا كانت الحروف التي تضاف إلى جذع الكلمة الإنجليزية تقتصر على الصدور والكواسع ، ولا تمس حروف الجذع ، فإن ما يضاف إلى جذع الكلمة العربية يشمل بالإضافة إلى الصدور والكواسع الحشو الذي يمكن أن يتخلل حروف الجذع فيغير معاملة . (المترجم) .

البيانات ، ثم يُعطى كل وجه وزناً رقمياً يحدده مسئول البحث تعسفياً ، كما يتم وضع «حد» معين . ويثبت الحد المتطلبات المنطقية لعملية البحث ، فهو الحد الأدنى للوزن الذي ينبغي أن تحققه تسجيلة الوثيقة لكي تسترجع . وأمامنا كمثال الاستراتيجية التالية :

$$\left\{ \begin{array}{l} \text{الجانب الخاص بالتقييم} \quad \circ \\ \text{الجانب الخاص بالحلقات النقاشية وبرامج التدريب} \quad \circ \\ \text{الجانب الخاص بعلم المعلومات} \quad \circ \end{array} \right. \text{الحد} = 15$$

فمسئول البحث يبحث عن الوثائق التي تتناول تقييم برامج التدريب والحلقات النقاشية في مجال علم المعلومات. وقد أعطى لكل جانب تعسفياً وزن ٥ ، وأقر الحد ١٥ . ويعني ذلك أن جميع الأوجه أو الجوانب الثلاثة ينبغي أن تكون ممثلة في تسجيلة الوثيقة لكي تسترجع . وهذا مناظر للاستراتيجية البولينية ١ و ب و ج ، حيث يمثل كل حرف وجهاً من أوجه الاستراتيجية . ومن الجدير بالملاحظة أن الوزن نفسه قد أعطي لجميع المصطلحات الخاصة بالوجه ، وأنه لا يمكن للوثيقة أن تحصل على وزن وجه معين إلا مرة واحدة فقط ؛ فمن الممكن للوثيقة في هذا المثال الحصول على خمس نقاط فقط لاشتمالها على الوجه الخاص بالتقييم ، مهما بلغ عدد مصطلحات التقييم التي حصلت عليها في الكشف . ومن الممكن استعمال الاستبعاد المنطقي (فيما عدا) باستعمال أوزان سالبة لأوجه الاستراتيجية .

وللبحث بالمصطلحات الموزونة مزية واحدة بارزة ، بالمقارنة بالأسلوب البوليني التقليدي ؛ فهو ييسر الحصول على نتائج البحث « مرتبة طبقياً » وفقاً لتسلسل أهميتها المحتملة بالنسبة للمستفيد . وأمامنا الطلب التالي على سبيل المثال :

أنا مهتم بالنشوي amyloidosis كأحد مضاعفات السل ، ويهمني بوجه خاص النشوي الكلوي renal ، كما أهتم في المقام الأول باستعمال البردنيوزون في علاج هذه الحالة .

ويتضمن هذا الطلب أربعة أوجه : (١) النشوي ، (٢) السل ، (٣) الكلى ، (٤) البردنيوزون . فإذا أعطينا هذه الأوجه الأوزان التالية :

| | |
|-------------|---|
| النشوي | ٥ |
| السل | ٥ |
| الكلبي | ٢ |
| البردنيوزون | ١ |

ووضعنا الحد عند « ١٠ » ، فإنه من الممكن الحصول على ناتج مرتب في أربعة مستويات:

- ١ - الوثائق التي أحرزت « ١٣ » (وهي التي يرجى منها تناول استعمال البردنيوزون في علاج النشوي الكلوي الناتج عن السل) .
 - ٢ - الوثائق التي أحرزت « ١٢ » (الوثائق التي تتناول النشوي الكلوي كأحد مضاعفات السل) .
 - ٣ - الوثائق التي أحرزت « ١١ » (استعمال البردنيوزون في علاج النشوي الناتج عن السل) .
 - ٤ - الوثائق التي أحرزت « ١٠ » (النشوي الناتج عن السل) .
- وتطبع تسجيلات الوثائق بهذا التسلسل ، وهو تسلسل تبعاً لتناقص الاتصال بطلب المستفيد . ولا تخرج الوثائق الواردة في المجموعة الرابعة عن حدود الطلب ، ومن ثم فإنها يمكن أن تكون متصلة إلى حد ما بطلب الحصول على المعلومات .
- هذا ، ومن الممكن الجمع بين الأسلوب البوليني وأسلوب المصطلحات الموزونة في بعض نظم التجهيز على دفعات . فأمامنا الاستراتيجية التالية على سبيل المثال :

| | |
|-----|----------|
| أ و | ب ١ (١٥) |
| | ب ٢ (١٤) |
| | ب ٣ (١٤) |
| | ب ٤ (١٣) |
| | ب ٥ (١٢) |

حيث تمثل الأرقام الواردة بين الأقواس الأوزان ، وتنص على أن الوثيقة ينبغي أن يكون بها مصطلح من « أ » ومصطلح من « ب » لكي تسترجع . إلا أن بعض مصطلحات « ب » أكثر أهمية من غيرها ، ولهذا فإن الوثيقة المكشوفة تحت « أ » و « ب ١ » تحت مرتبة أعلى من الوثيقة المكشوفة تحت « أ » و « ب ٢ » وهكذا . ومن الممكن التماس معالجة جيدة للبحث بالمصطلحات الموزونة في بحث سُمَار ودنس (Sommar and Dennis (1969) .

البحث الجزئي :

من الممكن أيضاً البحث في نظم الاسترجاع بإدخال خيط من المصطلحات يعبر عن الحاجة إلى المعلومات (من لغة مقيدة أو من كلمات نص حر) دون منطق رابط .
وحيث يقوم النظام بالبحث عن قطاعات من النص تضاهي الخيط الذي تم إدخاله على أحسن وجه . وعلى ذلك ، فإنه إذا كان الخيط الأصلي يتكون من خمس كلمات ، وهناك في مرصد البيانات وثيقة ما تشتمل على الكلمات الخمس مجتمعة ، فإن هذه الوثيقة تحصل على أعلى وزن ممكن ، وتتربع على قمة قائمة الوثائق المسترجعة .

ويعرف هذا الأسلوب بأسماء عدة : حيث يسمى بالبحث تبعاً لمستوى الربط ، والبحث الجزئي (Heaps and Sorenson (1968) والبحث تبعاً لدالة النصاب (Cleverdon, quorum function 1984) والاسترجاع وفق أفضل مضاهاة (Al-Hawamdeh et al. 1988) وهو نادراً ما يستخدم في الممارسة العملية ، على الرغم من أنه قد تبين بوجه عام أنه يسفر عن نتائج جيدة .

غربلة المخرجات :

يقوم اختصاصي المعلومات ، في بعض مراكز المعلومات ، بفحص المخرجات بعناية قبل تقديم النتائج للمستخدم ، حيث يستبعد الوثائق التي تبدو بجلاء غير متصلة بالموضوع ، وبذلك يرتفع بنسبة التحقيق في الناتج النهائي . ومما لا شك فيه أن نجاح عملية الغرلة هذه يرتبط مباشرة بمستوى بيان الطلب ، نظراً لأنه بناء على هذا الأساس يتم التنبؤ بالاتصال بالموضوع . فإذا كان بيان الطلب تعبيراً قاصراً عن حاجة المستخدم من المعلومات ، فإنه من الممكن لمسئول البحث أن يستبعد وثائق يمكن أن يراها المستخدم متصلة بالموضوع .

فالدقة إذن في بيان الطلب وفي تفسير مسئول البحث لاحتياجات المستخدم هي أهم العوامل المؤثرة في نجاح عملية الغرلة أو فشلها . وتشمل العوامل الأخرى مقدار الوقت المستنفذ في الغرلة ، ونوعية بدائل الوثائق التي يقدمها النظام . وكلما كان البديل مكتملاً كان من الأسر على مسئول البحث إصدار أحكام دقيقة على اتصال الوثائق المسترجعة بالموضوع ؛ فمن الممكن للعناوين بالإضافة إلى المصطلحات الكشفية أن تكون أكثر دلالة من العناوين وحدها ، كما يمكن للعناوين بالإضافة إلى المستخلصات

أن تكون أكثر دلالة من العناوين مع المصطلحات الكشفية . ويمكن للاعتماد على تسجيلات الوثائق في إصدار أحكام الاتصال بموضوع الطلب أن يكون مرتبطاً في الواقع، وعلى نحو مباشر بطول التسجيلة.

خصائص مسئولية البحث :

في دراسته لأثر الخبرة في البحث على الخط المباشر لم يجد فنيشل (Fenichel 1980 a, b; 1981) للعجب ، اختلافاً يذكر في السلوك ؛ فلم يكن مسئولو البحث المبتدئون أقل كفاءة من مسئولو البحث الأكثر تمرساً ، كما كان مسئولو البحث المتمرسون يستخدمون الأساليب البسيطة . كذلك لاحظ فنيشل تنوعاً كبيراً في سلوكيات البحث حتى بين المتساوين في مستوى الخبرة . كذلك تبين لهارتر (Harter 1984) من تحليل البيانات التي حصل عليها عن طريق الاستبانة وجود اختلافات شاسعة بين مسئولو البحث ، سواء في الاتجاه أو في السلوك . فقد كان الباحثون المتمرسون هم الأكثر حرصاً على استغلال الإمكانيات التفاعلية الكاملة لنظم الخط المباشر باتباع الأسلوب الإيعازي.

وقد قام كل من وانجر وماكدونالد وبرجر (Wanger, McDonald and Berger 1980) بدراسة ٥٣٥ عملية بحث أجراها على المدلين ١٩١ من اختصاصيي المعلومات، وتبين لهم أن عدداً قليلاً جداً من مسئولو البحث استغلوا الإمكانيات التفاعلية لنظام الخط المباشر ، وأن كفاءتهم (بناء على الاستدعاء والتحقيق) بدت غير مرتبطة لا بنوعية ما حصلوا عليه من تدريب ولا بدرجة تمرسهم في البحث في مرصد البيانات هذا . وكانت نتائج عمليات البحث ، في المتوسط ، محبطة إلى حد بعيد.

وقد أجريت دراسة ضافية للعوامل المؤثرة في فعالية عمليات البحث على الخط المباشر، على مدى عدة سنوات في مدرسة ماثيو باكستر للمكتبات وعلم المعلومات Mathew A. Baxter School of Library and Information Science التي لم يعد لها وجود الآن بجامعة كيس وسترن ريزيرف. وقد تناول ساراسفك وكانطور (Saracevic and Kantor 1988 a, b) وساراسفك وزملاؤه (Saracevic et al. 1988) المنهج ونتائج المرحلة النهائية لهذا المشروع. وكانت الدراسة تتناول ٤٠ مستفيداً ، تقدم كل منهم بسؤال واحد ، و ٣٩ من مسئولو البحث (ثلاثة منهم أعضاء بفريق المشروع و ٣٦ من غير الأعضاء) وقد تم تسجيل المقابلات التي أجريت مع المستفيدين على أشرطة . وقد أجريت لكل سؤال تسع

عمليات بحث ، أربع منها بواسطة أعضاء الفريق وخمس بواسطة مسئول البحث من غير الأعضاء ، وبذلك أصبحت مجموعة عمليات البحث ٣٦٠ عملية مختلفة . كذلك أجريت تجربة متصلة بالموضوع ، تنطوي على تصنيف أسئلة المستفيدين بواسطة ٢١ محكماً . وكانت نتائج عمليات البحث تقيّم من جانب المستفيدين بناء على الاتصال بالموضوع وإمكانية الإفادة منها . ويجعل التنوع المذهل لما تجمع من بيانات في المشروع من التلخيص الفعال ضرباً من المستحيل . وربما كانت أهم نتيجة انتهت إليها الدراسة أن مختلف مسئول البحث كانت لهم تفسيراتهم المختلفة للأسئلة ، ومن ثم فإنهم كانوا يتبعون أساليب بحث مختلفة ويسترجعون وثائق مختلفة . هذا بالإضافة إلى أن كل واحد من مسئول البحث كان يميل للعثور على وثائق تتصل بالموضوع لم يعثر عليها آخرون ، على الرغم من أن فرصة الحكم باتصال الوثائق المسترجعة بالموضوع كانت تزداد تبعاً لعدد من استرجعوها من مسئول البحث . ويرى الدارسون أن هذه النتائج تقدم دليلاً على الحاجة إلى تكرار عملية البحث بواسطة مسئول بحث مختلفين ، للسؤال نفسه ، إلا أن الخلاصة البديلة يمكن أن تكون هي احتمال تفضيل الطريقة الجماعية في تحليل السؤال ، والاتفاق الجماعي على استراتيجية البحث المبدئية ، كأسلوب لإجراء عمليات البحث ^(١) . وقد سبق أن تبين لفايدل (1985) Fidel أن مسئول البحث المتمرسين قلما يتفقون في اختيار المصطلحات ، كما كشفت الدراسات المبكرة أيضاً (بيتس (1977) Bates ولي (1954) Lilly على سبيل المثال) وبشكل مطرد عن أن المستفيدين من الفهارس البطاقية يميلون لعدم الاتفاق على أي المصطلحات يمكن أن تستخدم للبحث عن الوثائق المتخصصة في موضوع معين .

وينظر بلاكشو وفشوف (1988) Blackshaw and Fischhoff إلى البحث على الخط المباشر باعتباره عملية اتخاذ قرار . وقد تمت ملاحظة عدد من الباحثين المتطوعين أثناء إجراءات لعمليات بحث بالمؤلف أو العنوان أو الموضوع في فهرس متاح على الخط المباشر في مكتبة عامة . وسجل المؤلفان أن أداء مسئول البحث يشبه ذلك الذي كشفت عنه دراسات اتخاذ القرارات في السياقات الأخرى . كذلك تناول بلاردو (1985) Bellardo المتغيرات التي يمكن أن تؤثر في كفاءة الباحثين على الخط المباشر ، اعتماداً على مجموعتين اختباريتين من عمليات البحث التي أجراها طلبة ينتمون إلى ست من مدارس المكتبات

(١) تبين للانكستر وزملائه (1993) Lancaster et al. مؤخراً أنه من الممكن حتى لفريق من مسئول البحث المتمرسين أن يقدم نتائج محببة فعلاً .

المختلفة. وتدل النتائج على أنه من الممكن رد مظاهر الاختلاف في الأداء إلى الاستعداد اللفظي والكمي العام ، والإبداع الفني ، والميل نحو التفكير النقدي والتحليلي الإبداعي ، إلا أن جميع الارتباطات كانت واهية جداً . وقد لا يكون الذكاء المرتفع وغيره من العوامل التي غالباً ما يشار إليها باعتبارها من الخصائص المهمة ، من الأمور الضرورية للارتفاع بمستوى الأداء . ويذهب بلاردو إلى القول بأنه قد لا يكون من الممكن التنبؤ بكفاءة البحث بناء على السمات المعرفية وتلك الخاصة بالشخصية .

وتميز فايدل (1984) Fidel بين « مستولي البحث العمليين أو الإجرائيين Operationalist » و « مستولي البحث النظريين Conceptualist » ؛ فالأولون يميلون إلى اختيار أنسب المصطلحات والربط بينها ، وبذلك يتحركون في اتجاه التحقيق المرتفع ، بينما يميل الآخرون لتجميع قوائم شاملة بالمصطلحات الخاصة بكل وجه من أوجه البحث ، وبذلك يتجهون صوب الاستدعاء المرتفع.

وقد جرت خلال الخمسة والعشرين عاماً الماضية عدة محاولات لقياس جودة ما تقدمه المكتبات من خدمات مرجعية ، بطريقة بعيدة عن التطفل . وقد ركزت مثل هذه الدراسات على قدرة المكتبات على إجابة أسئلة الحقائق . ولماكيو (1988) McCue فضل الريادة في تقييم البحث على الخط المباشر في المكتبات بطريقة بعيدة عن التطفل . فقد طلب من إحدى وعشرين مكتبة عامة في شتى أنحاء الولايات المتحدة إجراء البحث نفسه في مرصدي بيانات مختلفين ، ولم يكن القائمون بإجراء عمليات البحث يعلمون أنهم تحت الاختبار . وقام خبيران محايدان في البحث في مرصد البيانات بتقييم النتائج . واعتماداً على نظام لتسجيل النقاط كانت كل عملية بحث تحصل على رصيدها من النقاط بناء على الاستراتيجية والأساليب العامة والنتائج . وكانت الأساليب العامة تشمل جوانب كأخطاء مسئول البحث في إدخال المصطلحات ، أو اتباع الإجراءات غير الصحيحة ، كذلك كانت النتائج تعطى أرصدة رقمية من النقاط ؛ فالإشارة الوراقية التي لا قيمة لها لا تحصل على نقاط بينما تحصل الإشارة الوراقية الممتازة على ثمانين نقاط . وكانت أرصدة المكتبات من النقاط تتراوح بين ١٥٥ و ٤١٩ . وانتهت ماكيو من تحليلاتها الإحصائية إلى أن المتغير الوحيد الذي ارتبط إيجابياً مع أرصدة النقاط المرتفعة لعمليات البحث هو عدد المواد المسترجعة . ولا عجب في ذلك على الإطلاق ؛ فواقع الأمر أن طريقة حساب النقاط تضع الاستدعاء في الاعتبار ، نظراً لأنه يعطي

أرصدة إيجابية للمواد « المفيدة » ، ولا تضع التحقيق في الاعتبار نظراً لأنه يعطي أصفاراً لا قيمة سلبية للمواد التي لا قيمة لها . إلا أن طريقة حساب الأرصدة برمتها مشكوك فيها ؛ فالنتائج هي العامل الحاسم على المدى الطويل ، ولهذا فإنه يبدو من العبث وضع طريقة لحساب النقاط التي يتم إحرازها ، لا تضع في اعتبارها النتائج فحسب ، وإنما تضع أيضاً استراتيجيات البحث وأساليب البحث .

وتقارن هانسن (1986) Hansen نتائج مجموعة من الأفراد غير المتمرسين قاموا بإجراء عمليات البحث يدوياً وأتبعوها بإجراء عمليات البحث على الخط المباشر ، بالنتائج التي حققتها مجموعة من المراقبة من الأفراد الذين أجروا عمليات البحث على الخط المباشر ، وأتبعوها بإجراء عمليات بحث يدوية . ولم يكن لأوجه الاختلاف في النتائج دلالة إحصائية . ورغم ذلك فقد انتهت هانسن إلى القول بأن الباحثين غير المتمرسين يمكنهم الحصول على أفضل نتائج عندما يجرون البحث يدوياً بعد إجراء البحث على الخط المباشر ، وذلك ، على الأقل في حالة « الموضوعات الضخمة المعقدة ذات الإنتاج الفكري المتشعب » وهي خلاصة لا تحظى إلا بدعم ضعيف جداً مما توافر لها من بيانات .

وتناقش فايدل (1991) Fidel كيف يمكن الاستفادة من تحليل سلوكيات البحث الخاصة بالوسطاء من البشر ، للتحقق من القواعد التي يختار بها الخبراء « مفاتيح البحث » ، أي الوصفات من اللغات المقيدة أو مصطلحات النصوص الحرة . وتدعي فايدل أنه مع إجراء المزيد من البحوث حول سلوكيات الخبراء في البحث ، سيكون من الممكن بناء قاعدة معرفية يمكن أن تستخدم في النظام الخبير الوسيط للبحث على الخط المباشر . ويقدم هوكنز (1988) Hawkins مراجعة علمية مفيدة للتطورات الحديثة في أساليب الذكاء الاصطناعي والنظم الخبيرة ، الخاصة بالبحث على الخط المباشر .

وعلى الرغم من الدراسة المكثفة لعملية البحث على الخط المباشر ، طوال السنوات العشرين الماضية ، تنتهي لوجان (1990) Logan إلى القول بأننا لم نتعلم بعد بما فيه الكفاية كيف نفسر الاختلافات الفردية في سلوكيات مستولي البحث .

العوامل المؤثرة في نجاح عملية بحث بعينها :

تعرضنا في هذا الفصل لمختلف العوامل المؤثرة في نجاح عملية البحث ، ونتناول تلك العوامل وثيقة الصلة ، وبشكل مباشر ، بعملية البحث ، بمزيد من التفصيل في الفصل العاشر . وهذه يمكن تلخيصها على النحو التالي :

١ - تفسير مسئل البحث لاحتياجات المستفيد : إن العامل الحاسم المؤثر في ناتج البحث هو نوعية التفاعل بين المستفيد والنظام . فإذا حدث أن كان بيان الطلب لا يمثل الحاجة إلى المعلومات بشكل ملائم ، فإنه ليس بإمكان مسئل البحث أن يفعل شيئاً لكي يقدم نتائج بحث جيدة ، إلا عرضاً ، أو رمية من غير رام .

٢ - مدى تعقد الطلب : فكلما كان الطلب بسيطاً ، أي ينطوي على عدد قليل من الأوجه ، كان من الممكن الحصول على نتائج جيدة للبحث . فالبحث الذي يطلب كل شيء تقريباً عن مرض تكهف النخاع Syringomyelia أحادي البعد ، وربما كان ينطوي على مصطلح كشفي واحد . وبمثل هذا الطلب العريض ، وعلى افتراض وجود مصطلح ملائم ، فإنه ينبغي أن يكون من الممكن الحصول على استدعاء مرتفع وتحقيق مرتفع . ولما كانت احتياجات المستفيد عامة ، فإنه يميل لقبول أية مقالة تتعرض بشكل واضح لموضوع تكهف النخاع ، حيث يعتبرها متصلة بالموضوع . وفي مقابل ذلك دعنا ننظر في طلب حول التغيرات الإشعاعية roentgenologic المشتركة في تكهف النخاع . فهذا طلب أكثر تعقداً ينطوي على ثلاثة أوجه أو أبعاد ، ويرتبط بعدد كبير من المصطلحات الكشفية ، فضلاً عن أهمية ما بين هذه المصطلحات من علاقات . هذا بالإضافة إلى أن معايير المستفيد في الحكم على الاتصال بالموضوع يمكن أن تكون أكثر صرامة ، حيث يمكن أن يرفض أية مقالة لا تتناول الموضوع الدقيق للتغيرات الإشعاعية في تكهف النخاع . وهناك طلب ثالث خاص بالانزلاق التلقائي للفقرة العنقية الأولى المحاكي لتكهف النخاع ، وهو طلب أكثر تعقداً أيضاً ، ينطوي على العلاقات المحددة بين المصطلحات الكشفية ، ويمكن للمستفيد أن يكون صارماً جداً في معايير الخاصة بالاتصال بالموضوع . ومع مثل هذا النوع من الطلبات تظهر أيضاً احتمالات العلاقات الغامضة أو الزائفة بين المصطلحات . ومن الممكن تسجيل استدعاء مرتفع لأي من عمليات البحث الثلاث هذه ، إلا أن ذلك لن يتم إلا عند مستوى تحقيق يمكن أن يتناقص بشكل ملحوظ تبعاً لتعقد الطلب .

٣ - قدرة لغة التكشيف على التعبير الدقيق عن الموضوعات : ينبغي أن تكون لغة النظام قادرة على التعبير عن موضوع الاستفسار بمستوى مناسب من التخصيص . كذلك ينبغي أن تكون اللغة مشتملة في بنائها على العلاقات التفريعية والعلاقات

الارتباطية ، بما يساعد مسئول البحث في صياغة الاستراتيجية . وينبغي أن تساعد بوجه خاص في إجراء عمليات البحث العامة بتوجيه الانتباه نحو جميع المصطلحات اللازمة لتحقيق التغطية الشاملة .

٤ - المجال الموضوعي الدقيق للطلب : من الممكن أن يكون هناك في أي مركز للمعلومات ، مجالات موضوعية معينة ، يمكن لأداء مسئول البحث أن يكون فيها أسوأ بوجه عام ، عما هو عليه في المجالات الموضوعية الأخرى . ومن الممكن بالنسبة لهذه الموضوعات أن تكون اللغة غير دقيقة ، أو تكون لغة مرصد البيانات ضعيفة ، أو يكون من الممكن للعلاقات الغامضة أو الزائفة بين المصطلحات أن تكون أكثر مما هي عليه في المجالات الأخرى .

٥ - سياسات التكشيف وممارساته : فإذا كان لدينا طلب للمقالات التي تتناول أخذ عينات من الخصيتين لدراساتها مجهرياً في حالات عدم الإخصاب ، فإنه من الممكن الحصول على استدعاء مرتفع لهذا الطلب ، وبنسبة تحقيق مقبولة على الأقل ، إذا كان من سياسة التكشيف استعمال مصطلح عينة مجهرية BIOPSY حيثما يرد بالمقالة ذكر لإجراء التحليل المجهري للعينة . وكذلك الحال أيضاً ، لا يمكن إجراء بحث شامل عن جميع تطبيقات أسلوب بولهاوزن Pohlhausen ، إذا كان منتج مرصد البيانات ، بناء على سياسة معينة ، لا يكشف الأساليب الرياضية عندما يرد ذكرها في الوثائق .

٦ - قدرات برمجيات البحث: ويحكم هذا الجانب ما تستطيع عملية البحث أو لا تستطيع القيام به على وجه التحديد؛ ما إذا كانت قادرة على بتر المصطلحات مثلاً ، وإذا كانت فعلاً فآين ؟

٧ - ملائمة الاستراتيجية ودقتها : تتوقف جودة الاستراتيجية على ما يلي :

أ - قدرة مسئول البحث على اختيار المصطلحات الصحيحة للتعبير عن موضوع البحث.

ب - قدرة مسئول البحث على وضع المصطلحات معاً ، على نحو سليم منطقياً.

ج - قدرة مسئول البحث على تدبر جميع المسالك المناسبة للاسترجاع .

د - قدرة مسئول البحث على صياغة الاستراتيجية، بتعديل الشمول أو التخصيص، بما يناسب متطلبات الاستدعاء والتحقيق ، أو قدرات احتمال المستفيد.

الفصل التاسع

معايير التقييم وإجراءاته

عادة ما يقيّم المستفيدون من الخدمات أيًا كان نوعها ، هذه الخدمات ، بوعي أو دون وعي ، بناء على التكلفة والوقت ومدى الجودة . فالسائح الذي يخطط للطيران من ميامي إلى بوينس آيرس ، ربما يسأل أول ما يسأل ما إذا كانت إحدى شركات الطيران العاملة على هذا الخط أقل تكلفة من غيرها . فإذا لم يكن الأمر كذلك أو كانت الفروق ضئيلة لا تذكر ، فإن عامل الوقت ربما يصبح أهم من كل ما عداه ، حيث يبحث المسافر عن أسرع رحلة ، أو الرحلة التي تغادر المطار في أنسب وقت . وإذا تساوت عدة رحلات في عامل الوقت يبدأ إدراك المسافر لنوعية شركة الطيران ، بناء على خبرته الشخصية أو خبرة الأصدقاء ، يؤثر في قراره .

المعايير المستخدمة في تقييم نظم استرجاع المعلومات :

يميل المستفيدون من نظم استرجاع المعلومات أيضاً للحكم عليها بمعايير التكلفة والوقت ومدى الجودة . والتكلفة مهمة في تقييم نظم استرجاع المعلومات ، كما هي مهمة في تقييم الخدمات والمنتجات الأخرى . فينبغي أن تقدم الخدمة بالتكلفة أو بالمقابل الذي يشعر المستفيد أنه معقول بالنسبة لعائده . وتنطوي التكلفة بالنسبة للمستفيد على ما هو أكثر من الرسوم المباشرة ؛ فهي تشمل تكلفة وقت المستفيد ، أي ما مقدار الجهد المستنفد في الاستفادة من النظام ، سواء من جانب المستفيد أو في تفويض عملية البحث إلى وسيط . وقد تبين من دراسات سلوكيات البحث عن المعلومات الخاصة بالعلماء وغيرهم

من المهنيين ، وبشكل مطرد ، أن قرب المثال وسهولة الإفادة هما العاملان الرئيسان المؤثران في اختيار مصدر المعلومات (Allen and Gerstberger, 1966, Rosenberg, 1966) فعادة ما يكون الاختيار لأقرب المصادر مثالاً ، سواء كان أو لم يكن في نظر المستفيد هو الأشمل أو الأوثق أو « الأفضل » . وتشمل عوامل سهولة الإفادة سهولة استجواب النظام، أي التعريف بحاجة المستفيد ، وسهولة الإفادة من المخرجات ، وخاصة للحكم على اتصال الوثائق المشار إليها بموضوع الحاجة إلى المعلومات . وهناك جانب مهم جداً لهذا المعيار الأخير ، وهو مدى توافر الإمداد الكفء الميسر بالوثائق . فالخدمة التي تقتصر على مجرد الإمداد بالإشارات الوراقية لا تقطع سوى شوط واحد على الطريق المفضي إلى تلبية احتياجات المستفيد من المعلومات ، ويمكن أن تكون سبباً في قدر كبير من الإحباط ، إذا لم يستطع المستفيد الحصول على الوثائق المشار إليها ، أو لم يكن باستطاعته الحصول على هذه الوثائق إلا من خلال إجراءات يراها مرهقة ومبذرة للوقت.

وللمستفيدين من خدمات المعلومات أنواع متعددة من الحاجة إلى المعلومات ، وتشمل :

١ - وثيقة بعينها معروفة الهوية .

٢ - معلومات حقائقية معينة يمكن أن ترد في نوع ما من الكتب المرجعية ، أو في أحد بنوك المعلومات الإلكترونية ، كالخواص الفيزيائية الحرارية لمادة معينة مثلاً.

٣ - عدد قليل من المقالات « الجيدة » في موضوع معين ، أو الإشارات الوراقية الخاصة بهذه المقالات.

٤ - بحث شامل للإنتاج الفكري في مجال موضوعي معين.

٥ - خدمة إحاطة جارية تجعل المستفيدين على دراية بالإنتاج الفكري الحديث المتصل باهتماماتهم المهنية الجارية.

ولهذه الخدمات متطلباتها المختلفة لوقت الاستجابة ؛ فخدمة الإحاطة الجارية مثلاً ينبغي أن تقدم بانتظام وعلى فترات متقاربة ، كما أن ما تقدمه من معلومات ينبغي أن يكون حديثاً قدر الإمكان . أما المستفيد الذي يحتاج إلى بحث شامل للإنتاج الفكري

فعادة ما يكون مشغولاً بمشروع بحث طويل المدى . ويمكن لسرعة الاستجابة ألا تكون عاملاً حاسماً إلا إذا كان هناك تاريخ معين يمكن بعده للنتائج أن تتضاءل أهميتها أو تفقد هذه الأهمية . وغالباً ما يكون المستفيد على استعداد للانتظار بعض الوقت للحصول على بحث شامل ، حيث اكتمال النتائج عنده أهم من السرعة . إلا أنه بالنسبة للأنواع الأخرى من الحاجة إلى المعلومات ، عادة ما يكون المستفيد حريصاً على الاستجابة السريعة .

ومعايير التكلفة والوقت المتصلة بتقييم نظم استرجاع المعلومات مباشرة إلى حد بعيد، كما أنها ثابتة نسبياً ولا تختلف كثيراً من نشاط إلى آخر . أما معايير الجودة فربما كانت أقل تحديداً أو أقل وضوحاً في معالها ، كما أنها تختلف وبشكل ملحوظ تبعاً للنظام الذي يتم تقييمه وتبعاً لاحتياجات المستفيدين .

وهناك مقياسان نوعيان رئيسان لنجاح نظم استرجاع المعلومات :

(١) هل يحصل المستفيد على ما يبحث عنه فعلاً ؟

(٢) ما مدى اكتمال ودقة ما يحصل عليه ؟

وأول هذين المقياسين ، والذي يمكن أن يطبق على البحث عن وثيقة بعينها ، أو على الإجابة عن سؤال عن حقائق معينة ، بسيط ولا لبس فيه على الإطلاق . أما المقياس الثاني فأكثر صعوبة إلى حد بعيد في التطبيق العملي ، نظراً لأنه ينطوي على كل من الأحكام القيمية من جانب البشر واستخدام نوع ما من أدوات القياس المتدرجة للتعبير عن درجات النجاح . ومع ذلك فإنه لا غنى عنه في تقييم معظم أنواع أنشطة استرجاع المعلومات . والاستدعاء والتحقيق اللذان عرضنا لهما في الفصل الثالث معياران غالباً ما يستخدمان في الحكم على كفاءة البحث في نظم استرجاع المعلومات .

(١) تقييم الفعالية .

(٢) تقييم فعالية التكلفة .

(٣) تقييم علاقة العائد بالتكلفة .

مستويات التقييم :

هناك ثلاثة مستويات محتملة يمكن عندها تقييم خدمات المعلومات (الشكل رقم ٣٨):

المستوى الأول : تقييم الفعالية (اعتبارات رضا المستفيد)

أ. معايير التكلفة :

١. التكلفة المالية بالنسبة للمستفيد (لكل عملية بحث أو لكل اشتراك أو لكل وثيقة).

٢. اعتبارات التكلفة الأخرى الأقل قابلية للتحديد الدقيق.

أ. الجهد المستنفد في تعلم كيفية الإفادة من النظام .

ب. الجهد المستنفد في الإفادة الفعلية .

ج. الجهد المستنفد في استرجاع الوثائق (عن طريق نظم الإمداد بالوثائق

والمساندة للنظام) .

د. شكل المخرجات التي يقدمها النظام .

ب. معايير الوقت :

١. الوقت ما بين تقديم الطلب واسترجاع الإشارات الوراقية .

٢. الوقت ما بين تقديم الطلب واسترجاع الوثائق .

٣. اعتبارات الوقت الأخرى ، كوقت الانتظار لاستخدام نظام الخط المباشر مثلاً .

ج. اعتبارات الجودة :

١. تغطية مرصد البيانات

٢. اكتمال المخرجات (الاستدعاء)

٣. اتصال المخرجات بموضوع الطلب (التحقيق)

٤. حداثة المخرجات

٥. اكتمال البيانات ودقتها

المستوى الثاني : تقييم فعالية التكلفة (رضا المستفيد بالنسبة للكفاءة الداخلية للنظام واعتبارات التكلفة)

أ. تكلفة الوحدة لكل إشارة وراقية متصلة بالموضوع تم استرجاعها

ب. تكلفة الوحدة لكل إشارة وراقية جديدة ، لم تكن معروفة من قبل ، وتتصل

بالموضوع وتم استرجاعها

ج. تكلفة الوحدة لكل وثيقة متصلة بالموضوع تم استرجاعها.

المستوى الثالث : تقييم عائد التكلفة (أهمية النظام في مقابل تكاليف تشغيله أو الإفادة منه).

الشكل رقم (٣٨) : المعايير التي يمكن بها تقييم نظم استرجاع المعلومات .

وتقييم فعالية خدمة المعلومات ، في ظروفه المثالية ، دراسة مدى قدرة الخدمة على تلبية احتياجات المستفيدين منها . وهو في الواقع أقرب ما يكون إلى دراسة مدى قدرة الخدمة على تلبية الطلبات ، أي الاحتياجات المعرب عنها من جانب المستفيدين ، أما الاحتياجات التي لم يتم الإعراب عنها ، أي الاحتياجات الكامنة أو المستترة للمستفيدين الحاليين ، واحتياجات غير المستفيدين ، فالتجاهل من نصيبها في غالب الأحيان . وتشمل أهم المعايير التي يمكن بها تقييم الفعالية ، التكلفة والوقت والجودة . وتشمل معايير الجودة التغطية ، ومدى اكتمال المخرجات (الاستدعاء) ومدى اتصالها بموضوع الطلب (التحقيق) ، ومدى حداثة البيانات واكتمالها ودقتها .

ويربط تقييم فعالية التكلفة بين مقاييس الفعالية ومقاييس التكلفة . فمن الممكن على سبيل المثال التحقق من مختلف الطرق التي يمكن بها الارتفاع بقدرة المكتبة على الإمداد بالوثائق من ٧٢٪ إلى ٨٠٪ . والتحليل الذي يبين أي هذه الطرق هو الأقل تكلفة ، تحليل لفعالية التكلفة .

وتربط دراسة عائد التكلفة تكاليف تقديم خدمة ما بعائدات توافر هذه الخدمة . وفي مجال تجهيز المعلومات عادة ما يكون تحليل عائد التكلفة صعباً نظراً للمشكلات التي تنطوي عليها محاولة تحديد القيمة المالية الفعلية للمعلومات .

الاستدعاء والتحقيق :

الاستدعاء مقياس لما إذا كان قد تم استرجاع وثيقة بعينها أم لا ، أو مقياس مدى استرجاع الوثائق المراد استرجاعها . وفي حالة المستفيد الذي يلتبس وثيقة بعينها ، فإن الوثيقة إما أن تسترجع من المجموعة ، أي تستدعى في الوقت الذي تدعو الحاجة إليها فيه ، أو لا تسترجع . أما في حالة المستفيد الذي يرغب في إجراء بحث شامل لمرصد البيانات ، فإنه من الممكن التعبير عن نجاح البحث بناء على مدى استرجاع الوثائق المتصلة بموضوع الاهتمام أو استرجاع الإشارات الوراقية الخاصة بهذه الوثائق . وغالباً ما يسمى مقياس اكتمال البحث في مرصد البيانات بنسبة الاستدعاء ، وتعني عبارة « ٨٠ بالمئة استدعاء » أنه قد تم العثور على أربعة أخماس الوثائق المتصلة بموضوع الاهتمام في مرصد البيانات .

ويدل التحقيق على مقياس لمعدل الإشارات إلى الشوشرة في أنواع معينة لنظم المعلومات .
فبحث الإنتاج الفكري الذي يسترجع ٥٠ وثيقة ، حُكم على عشر منها من جانب طالب البحث
بأنها متصلة بموضوع الاهتمام ، يمكن أن يقال إنه يعمل بمعدل تحقيق ١٠ / ٥٠ أو ٢٠
بالمئة .

ومن الممكن الاستطراد في وصف نسب الاستدعاء والتحقيق بجدول ٢×٢ يشتمل على
النتائج التي تحققت عند إجراء بحث معين للإنتاج الفكري (الشكل رقم ٣٩) . فعندما يتم
إجراء البحث في معظم نظم استرجاع المعلومات ، يقسم النظام المجموعة إلى جزئين: فالوثائق
التي تضاهاى استراتيجية البحث التي استخدمت في استجواب النظام يتم استرجاعها (أ + ب)
والوثائق التي لا تضاهاى الاستراتيجية لا يتم استرجاعها (ج + د) . ومن الممكن النظر إلى هذا
التقسيم الثنائي لمجموعة الوثائق باعتباره شكلاً من أشكال قدرة النظام على التنبؤ بالوثائق
المتصلة بالموضوع ؛ فالنظام يتنبأ ، على نحو ما ، بأن هناك وثائق معينة يمكن أن تكون
متصلة بموضوع البحث وأخرى قد لا تكون متصلة ، فيسترجع الأولى ويترك الأخيرة .

| تنبؤ النظام بالاتصال بالموضوع | قرارات المستفيد الخاصة بالاتصال بالموضوع | | |
|-------------------------------|--|-----------------------------|--------------------------------------|
| | متصل بالموضوع | غير متصل بالموضوع | المجموع |
| مسترجع | ١ إصابات | ب (شوشرة) | ١ + ب |
| غير مسترجع | فاقد ج | (غير صالح للاسترجاع) د | ج + د |
| المجموع | ١ + ج | ب + د | ١ + ب + ج + د (إجمالي المجموعة) |

الشكل رقم (٣٩) : جدول ٢ × ٢ لنتائج البحث في الإنتاج الفكري .

وفي جميع المواقف تقريباً ، دائماً ما يكون عدد الوثائق المسترجعة بالبحث قليلاً
جداً بالنسبة لإجمالي حجم المجموعة . وبعبارة أخرى عادة ما يكون (أ + ب) صغيراً
أمّا (ج + د) أي عدد الوثائق التي لم تسترجع فكثير جداً ؛ فمن الممكن لبحث
ما أن يسترجع ٨٠ تسجيلية من ملف يشتمل على ٥٠٠,٠٠٠ تسجيلية . وفي هذه
الحالة أ + ب = ٨٠ و ج + د = ٤٩٩,٩٢٠ .

ويرتبط البعد الآخر للجدول 2×2 بقرارات الاتصال بالموضوع التي يتخذها من طلب إجراء البحث؛ فالبحث الكامل يسترجع جميع الوثائق الواردة في مرصد البيانات والتي يحكم الاستفادة باتصالها بالموضوع (أ + جـ) ولا يسترجع أيًا من تلك التي يحكم الاستفادة بعدم اتصالها بالموضوع (ب + د). وهناك تطابق تام بين قرارات الاتصال بالموضوع التي يتخذها الاستفادة وتوقع النظام لما يتصل بالموضوع، أي أن ب = صفراً و جـ = صفراً. وقد حققت عملية البحث هذه ١٠٠ بالمئة استدعاء، كما بلغت أيضاً ١٠٠ بالمئة تحقيقاً.

ويرتبط الاستدعاء إذن بقدرة النظام على استرجاع الوثائق المتصلة بالموضوع، بينما يرتبط التحقيق بالقدرة على عدم استرجاع الوثائق غير المتصلة بالموضوع. ومن الممكن التعبير عن درجات الاستدعاء والتحقيق التي يتم إحرازها في بحث ما كنسب؛ فنسبة الاستدعاء تحسب على النحو التالي:

$$100 \times \frac{\text{عدد الوثائق المسترجعة المتصلة بالموضوع}}{\text{مجموع الوثائق المتصلة بالموضوع في المجموعة}}$$

وبناء على الشكل رقم (٣٩) فإن نسبة الاستدعاء هي (أ / [أ + جـ]). أما نسبة التحقيق فتحسب على النحو التالي:

$$100 \times \frac{\text{عدد الوثائق المسترجعة المتصلة بالموضوع}}{\text{مجموع الوثائق المسترجعة}}$$

وبناء على الشكل رقم (٣٩) فإن نسبة التحقيق هي (أ / [أ + ب]).

وتعبر نسب كل من الاستدعاء والتحقيق عن قدرة النظام على التنقية، أي قدرته على استرجاع المطلوب وحجب ما سواه. ولا يمكن لأي من الاستدعاء أو التحقيق منفرداً أن يعطي صورة كاملة لفعالية البحث. ومن الممكن دائماً الحصول على ١٠٠ بالمئة استدعاء؛ فإذا استرجعت المجموعة (أ + ب + جـ + د)، فإن الاستدعاء يكون ١٠٠ بالمئة. إلا أنه يمكن للأسف أن يكون التحقيق متدنياً إلى أبعد حد، نظراً لأنه بالنسبة لأي طلب عادة ما تكون الغالبية العظمى من الوثائق التي تضمها المجموعة غير متصلة بالموضوع.

ويمكن النظر إلى نسبة التحقيق باعتبارها نوعاً من عوامل التكلفة في وقت المستفيد؛ أي الوقت اللازم للتحقق من الإشارات الوراقية المتصلة بالموضوع في مخرجات عملية البحث. فإذا نظرنا إلى طلب للبحث يتوافر له بأحد مراصد البيانات عشرون وثيقة متصلة بالموضوع، وافترضنا اتباع ثلاث استراتيجيات بحث مختلفة لاستجواب النظام، وأن كل استراتيجية تسترجع ١٥ وثيقة من الوثائق العشرين المتصلة بالموضوع، أي بنسبة استدعاء ٧٥ بالمئة. وبلاستراتيجية الأولى تم استرجاع ٣٠ وثيقة، وبالثانية ٦٠ وثيقة، وبالثالثة ١٥٠ وثيقة، أي أن نسب التحقيق في هذه الاستراتيجيات الثلاث ٥٠ بالمئة و ٢٥ بالمئة و ١٠ بالمئة على التوالي. وكان على المستفيد في الاستراتيجية الأولى أن يتفحص ٣٠ تسجيلة فقط ليعثر على الخمس عشرة المتصلة بالموضوع، وأن يتفحص في الثانية ٦٠ تسجيلة وفي الثالثة ١٥٠ تسجيلة. وإذا ما تساوت جميع عناصر الموقف، فإن المستفيد يستنفد في التحقق من التسجيلات المتصلة بالموضوع، في الاستراتيجية الثانية أطول مما يستنفده في الاستراتيجية الأولى، كما يستنفد في التحقق من التسجيلات المتصلة بالموضوع في الاستراتيجية الثالثة وقتاً أطول بكثير جداً مما يستنفده في الأولى أو الثانية. وعلى هذا النحو يمكن النظر إلى نسبة التحقيق كمقياس لجهد المستفيد أو ما يتحمله من تكلفة. فعملية البحث التي تحقق ٧٥ بالمئة استدعاء عند مستوى ٥٠ بالمئة تحقيق أكثر كفاءة من تلك التي تحقق ٧٥ بالمئة استدعاء عند مستوى ٢٥ بالمئة تحقيق، والتي بدورها أكثر كفاءة من تلك التي تحقق ٧٥ بالمئة استدعاء عند مستوى ١٠ بالمئة تحقيق.

وتقيس هذه النسب درجة التطابق بين قرارات المستفيد الخاصة بالاتصال بالموضوع وتنبؤات النظام للاتصال بالموضوع. وفي البحث الكامل تتطابق كل من القرارات والتنبؤات تماماً. وعمليات البحث الكاملة نادرة بالطبع إلى أبعد الحدود. والاحتمال الأرجح هو الموقف الذي يكون فيه تطابق جزئي بين المجموعة (أ + ج) والمجموعة (أ + ب) (الشكل رقم ٤٠). وقد استرجع هذا البحث الافتراضي، الذي لا يختلف كثيراً عن الواقع الفعلي، معظم وليس كل الوثائق المتصلة بالموضوع، كما تجنب معظم وليس كل الوثائق غير المتصلة بالموضوع.

ويميل الاستدعاء والتحقيق للارتباط العكسي؛ فعند توسعة عملية البحث لتعزيز الاستدعاء يميل التحقيق للانخفاض. وفي مقابل ذلك، فإنه عند تضيق مجال البحث

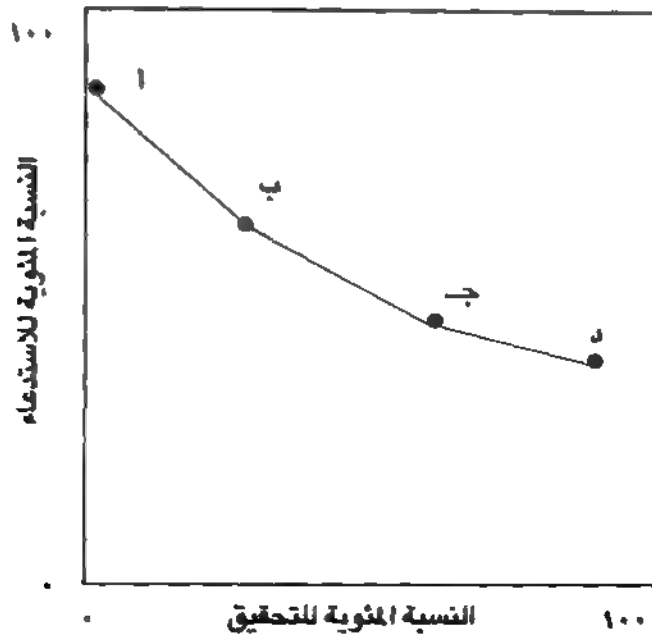
لتعزيز التحقيق يميل الاستدعاء للانخفاض . ومن الممكن بالنسبة لمجموعة معينة من الطلبات إجراء عمليات بحث لكل منها على عدة مستويات مختلفة ، تتراوح بين البحث العريض إلى أبعد حد ، والمخطط للحصول على استدعاء مرتفع ، والبحث الضيق إلى أبعد حد، والمخطط للحصول على تحقيق مرتفع . فإذا تم حساب نسب الاستدعاء والتحقيق الخاصة بكل عملية من عمليات البحث هذه ، وتوقيعها في مقابل بعضها البعض، فإن الرسم البياني يمكن أن يبدو قريب الشبه بالمنحنى الوارد في الشكل رقم (٤١) . ويمثل الرسم البياني متوسط نسب الاستدعاء والتحقيق الخاصة بكل عمليات البحث ، حيث أجريت كل عملية بحث على أربعة مستويات متفاوتة . وعندما تجري عمليات البحث على مستوى عريض جداً (النقطة أ) يتحقق استدعاء مرتفع جداً يصل إلى حوالي ٩٠ بالمئة، إلا أن التحقيق يبدو منخفضاً جداً . وعند إجراء عمليات البحث بشكل مخصص جداً فإن ذلك يؤدي إلى الحصول على تحقيق مرتفع واستدعاء منخفض (النقطة د) . وتمثل النقطتان ب و ج الاستراتيجيات التي تحاول الموازنة بين الطرفين، الاستغراق في التعميم والمبالغة في التخصيص .



الشكل رقم (٤١) : نتائج البحث في نظام لاسترجاع المعلومات . والمساحة المظلمة هي كل الوثائق المسترجعة، وغير المظلمة هي الوثائق التي لم تسترجع . وقد استرجع البحث معظم الوثائق المتصلة بالموضوع (أ) وفاته عدد قليل من الوثائق المتصلة بالموضوع (ج) واسترجع عدداً من الوثائق غير المتصلة بالموضوع (ب) وتقع غالبية المجموعة (أ + ب + ج + د) في فئة الوثائق التي أصاب النظام في رفضها (د) .

ولا يحتاج الكل دائماً إلى الاستدعاء المرتفع ، حيث تختلف متطلبات الاستدعاء والتحقيق باختلاف المستفيدين ، كما يمكن أن تختلف متطلبات الفرد الواحد من ظرف إلى آخر . ويمكن لتحمل التحقيق المنخفض أن يكون مرتبطاً وبشكل مباشر بمتطلبات الاستدعاء الخاصة بالمستفيد ؛ فعلى أحد الطرفين نجد الشخص الذي يؤلف كتاباً أو يعد مراجعة علمية أو يبدأ مشروع بحث طويل المدى . ويمكن لهذا الشخص أن يرغب في

البحث الشامل (استدعاء مرتفع) ومن ثم فإنه يمكن أن يتحمل التحقيق المنخفض إلى حد ما ، لكي يضمن عدم إفلات شيء ذي أهمية . وعلى الطرف الآخر نجد المستفيد العادي من أحد نظم استرجاع المعلومات الصناعية مثلاً ، والذي يحتاج إلى عدد قليل من المقالات الحديثة حول موضوع ما فوراً . فهو لا يحتاج إلى الاستدعاء المرتفع وإنما إلى التحقيق المرتفع . ويمكن للبعض الآخر أن يفضل المواءمة بين مستوى «معقول» للاستدعاء عند مستوى «مقبول» للتحقيق .



الشكل رقم (٤١) : الرسم البياني للاستدعاء مقابل التحقيق.

ومن العبث استخدام نسبة الاستدعاء في قياس نجاح عملية البحث عندما يكون الاستدعاء المرتفع لا أهمية له . وقد حدث هذه الحقيقة بالبعض إلى اقتراح مقياس للاستدعاء التناسبي أو النسبي ، يتم فيه التعبير عن نجاح عملية البحث بناء على عدد الوثائق المسترجعة المتصلة بالموضوع مقسوماً على عدد الوثائق المتصلة بالموضوع والتي يريدونها المستفيد . فإذا أعرب المستفيد ، على سبيل المثال ، عن حاجته إلى خمس وثائق متصلة بالموضوع ، إلا أن البحث لم يسترجع سوى ثلاث فقط ، حينئذ يكون الاستدعاء التناسبي $3 / 5$ ، أي ٦٠ بالمئة . وعلى الرغم من أن هذا المقياس يبدو جذاباً لأول وهلة ، فإنه منقطع الصلة بالواقع ، لأنه لا يستطيع تحديد عدد الوثائق التي يرغبها من النظام مسبقاً سوى عدد قليل جداً من المستفيدين .

وجه آخر من أوجه القصور في نسبة الاستدعاء ، وهو أنه يفترض بشكل أو بآخر أن جميع الوثائق المتصلة بالموضوع متساوية تقريباً في أهميتها . ولا يصدق ذلك دائماً؛ فمن الممكن لعملية بحث أن تسترجع خمس وثائق متصلة بالموضوع ويفلت منها عشر (نسبة الاستدعاء = ٣٣ بالمئة) إلا أن الوثائق الخمس التي استرجعت يمكن أن تكون أفضل بكثير من العشر التي أفلتت ، حيث يمكن أن تكون أحدث ، ويمكن أن تجعل في الواقع من الوثائق العشر الأخرى مجرد حشو لا قيمة له تقريباً . وعلى ذلك ، فإن نسبة الاستدعاء رغم أهميتها ، ينبغي أن تستخدم بحذر في تقييم نظم استرجاع المعلومات .

ولنسبة التحقيق أيضاً بعض أوجه القصور ؛ فهي فعلاً مقياس غير مباشر لما ينفقه المستفيد من وقت وجهد في مرحلة المخرجات الخاصة بعملية استرجاع المعلومات؛ أي أنه كلما ارتفعت نسبة التحقيق قل الجهد الذي يمكن أن يستنفده المستفيد في التحقق من الوثائق المتصلة بالموضوع . ففي بحث أسفر عن نسبة تحقيق منخفضة جداً ، استرجعت فيه مثلاً عشر وثائق فقط من بين ثمانين وثيقة تتصل بالموضوع ، فإن الأمر قد يتطلب قدراً كبيراً جداً من وقت المستفيد وجهده ، للتحقق من الوثائق المتصلة بالموضوع ، في القائمة المطبوعة ، وخاصة إذا كانت هذه القائمة لا تشتمل إلا على الإشارات الوراقية ، ويتعين على المستفيد استرجاع نسخ من عدد كبير من الوثائق حتى يقرر أيها يتصل بالموضوع وأيها لا يتصل . ويستخدم مقياس الجهد هذا في تقييم البحث المفوض ، أي البحث الذي يجريه أحد اختصاصيي المعلومات نيابة عن المستفيد . وفي هذا الموقف ينظر إلى النظام باعتباره « صندوقاً أسود » تقريباً ، يوضع فيه الطلب وتخرج منه مجموعة من الوثائق أو الإشارات الوراقية الخاصة بالوثائق . ونسبة التحقيق مقياس صالح لكفاءة أي نوع من عمليات البحث المفوض ، التي يقدم فيها المستفيد الطلب إلى النظام وينتظر النتائج ، سواء كان البحث يتم يدوياً أو إلكترونياً .

وتفقد نسبة التحقيق جزءاً من مغزاها عندما تطبق على البحث غير المفوض؛ فالمستفيد هنا يقوم بإجراء بحثه بنفسه ، ويتخذ قرارات الاتصال بالموضوع دون توقف أثناء إجراء البحث ، ذلك أن المستفيد عندما يرجع إلى التسجيلات الواردة تحت أحد المداخل الكشفية في أحد الكشافات المطبوعة أو في أحد نظم الخط المباشر ، فإنه يترك الوثائق غير المتصلة بالموضوع جانباً ولا يسجل سوى تلك التي تبدو متصلة بالموضوع . ومن الممكن تحديد نسبة التحقيق بالنسبة لهذا النوع من عمليات البحث

بحساب إجمالي عدد الإشارات الوراقية التي رجع إليها المستفيد ، وعدد تلك التسجيلات التي أقر اتصالها بالموضوع ، وقسمة الأخيرة على الأولى . أما جهد المستفيد في البحث غير المفوض فيمكن التعبير عنه بشكل أكثر مباشرة بناء على الوقت اللازم لإجراء البحث، الذي يمكن الخروج منه بتكلفة الوحدة (بالوقت) لكل وثيقة متصلة بالموضوع أمكن العثور عليها . ومن المفترض أنه كلما ارتفع التحقيق في البحث غير المفوض قل الوقت المستنفد في البحث ، إذا تساوت جميع عناصر الموقف الأخرى .

وفضلاً عن التكاليف المباشرة ، ناقشنا حتى الآن أربعة معايير للكفاءة ، يمكن بها تقييم أي نوع من أنواع البحث في الإنتاج الفكري ، يدوياً كان أو إلكترونياً ، وذلك من وجهة نظر رضا المستفيد ؛ وهذه المعايير هي الاستدعاء ، والتحقيق ، وزمن الاستجابة ، وجهد المستفيد . وفيما يلي أبرز نقاط مقاييس الكفاءة هذه .

١ - الاستدعاء : مهم بالنسبة للمستفيدين من نظم استرجاع المعلومات الذين يبحثون عن أوعية المعلومات في موضوع معين . وعادة ما يتطلب الأمر الحد الأدنى للاستدعاء لا أكثر ، أي كتاباً واحداً مثلاً أو عدداً قليلاً من المقالات حول موضوع بعينه . إلا أنه يمكن الحرص في بعض الأحيان على الحد الأقصى للاستدعاء . وأمامنا كمثال المستفيد الذي يريد إجراء بحث شامل في خدمة الاستخلاص *Chemical Abstracts* .

٢ - التحقيق : مقياس غير مباشر لوقت المستفيد وجهده ، وأقل صلاحية في تقييم عمليات البحث غير المفوضة في مرصد البيانات .

٣ - جهد المستفيد : يقاس الجهد ، في البحث غير المفوض بمقدار الوقت الذي يستنفده المستفيد في إجراء البحث . أما في البحث المفوض فإنه يقاس بمقدار ما يستنفده المستفيد من وقت في التفاوض حول الطلب مع النظام ، ومقدار الوقت اللازم ، بعد الحصول على نتائج البحث ، للتحقق من الوثائق المتصلة بالموضوع ، ويرتبط جهد المستفيد مباشرة بنسبة التحقيق .

٤ - زمن الاستجابة : وهو في البحث المفوض الوقت الفاصل بين تقديم الطلب من جانب المستفيد وتلقي نتائج البحث . أما في البحث غير المفوض فهو الوقت المستنفد في إجراء البحث فعلاً . ويعد في هذه الحالة مقياساً أيضاً لجهد المستفيد .

والارتباط وثيق بين كل هذه المعايير وبعضها البعض ، كما أنها تتبادل الأدوار فيما بينها؛ فالمستفيد الذي يريد استدعاءً مرتفعاً عادة ما يتحمل التحقيق المنخفض ، حيث يكون لديه الاستعداد لبذل مزيد من الجهد الشخصي في البحث ، وربما يسلم أيضاً ببعض التأخير في البحث . أما المستفيد الذي يتطلب الحد الأدنى من الاستدعاء فيمكن أن يتوقع تحقيقاً مرتفعاً وزمن استجابة سريعاً ، وربما لا يكون على استعداد لبذل جهد شخصي يذكر في البحث . ومن المهم أن ندرك أن زمن الاستجابة دائماً ما يكون ثانوياً بالنسبة للاستدعاء والتحقيق . وحتى عندما يكون زمن الاستجابة أمراً أساسياً، كما هو الحال مثلاً في مركز معلومات السموم ، فإن المطلب الأول هو أن تقدم المعلومات، وأن تكون متناهية الدقة ، أي أنه ينبغي أن يكون هناك بعض الاستدعاء، ونسبة تحقيق ١٠٠ بالمئة ، وعلى ذلك فإن زمن الاستجابة رغم أهميته البالغة يبدو ثانوياً. ومن العيب فعلاً وضع زمن الاستجابة في صدارة معايير التقييم ، لأن ذلك يمكن أن يعني أن المستفيدين يمكن أن يفضلوا الحصول الفوري على مصادر للمعلومات لا تتصل بالموضوع على الحصول المتأخر على المصادر المتصلة بالموضوع .

المقاييس الأخرى لتقييم عمليات البحث :

يشتمل الجدول ٢×٢ (انظر الشكل رقم ٢٩) على جميع البيانات التي تدعو الحاجة إليها لتقييم البحث في نظم الاسترجاع. ومن بين طرق التعبير عن هذه المعلومات حساب نسب الاستدعاء والتحقيق ، إلا أنه من الممكن اشتقاق عدة مقاييس أخرى من الجدول، والتعبير عنها بطرق مختلفة .

واعتماداً على الشكل رقم (٢٩) يمكن اشتقاق المقاييس التالية :

١ / (أ + جـ) نسبة الاستدعاء ، ويعرف أيضاً « بمعدل الإصابة » . وقد اقترح هذا المقياس لأول مرة من جانب آلن كنت وزملائه (Kent et al. (1955) ، حيث كان يسمى «عامل الاستدعاء» . كذلك أسماه سوتس (Swets (1963 « الاحتمال الشرطي للإصابة» . ويسميه جوفمان ونيول (Goffman and Newill (1964 «الحساسية» .

جـ / (أ + جـ) متم الاستدعاء . وقد أسماه فيرثورن (Fairthorne (1965 «نسبة الاقتناع» ، بينما يسميه سوتس (Swets (1963 « الاحتمال الشرطي للخطأ » .

أ / (أ + ب) نسبة التحقيق ، وتسمى أحياناً « بنسبة الاتصال بالموضوع » . وكنت وزملاؤه (Kent et al. (1955 هم أول من أدخل هذا المقياس ، حيث أسموه «عامل الصلاحية» . ويسميه آخرون « معدل القبول » .

ب / (أ + ب) متمم نسبة التحقيق، ويسمى أحياناً «عامل الضوضاء» (Kent et al. (1955 .
ب / (ب + د) وربما كان سوتس (Swets (1963 أول من اقترحه ، حيث أسماه «الاحتمال الشرطي للقطرات الزائفة» ، وأسماء كل من كلفردون وملزوكين (Cleverdon (Mills and Keen 1966 فيما بعد «نسبة السقط» . كما يسمى أيضاً « المستبعد » .

د / (ب + د) متمم السقط ، ويسميه جوفمان ونيول (Goffman and Newill (1964 «التخصيص» . أما سوتس (Swets (1963 فيسميه « الاحتمال الشرطي للرفض الصحيح » .

ويعبر كل مقياس من هذه المقاييس عن جانب واحد فقط من نتائج البحث . وعند استخدام مقياسين معاً ، كما في الرسم البياني للاستدعاء في مقابل التحقيق ، أو الاستدعاء في مقابل السقط ، فإنه يسمى « المقياس ثنائي المتغير » . وعند الجمع بين مقياسين من هذه المقاييس المنفردة في مقياس واحد ، يعبر على سبيل المثال عن كل من الاستدعاء والتحقيق ، فإن ذلك يسفر عن « مقياس مركب » أو « رقم جدارة موحد » .

وهذه المقاييس ملائمة لنظام الاسترجاع الذي يقتصر على تقسيم المجموعة إلى قسمين، تلك المواد التي تسترجع بعملية بحث معينة ، وتلك المواد التي لا تسترجع. إلا أن هناك بعض النظم التي تفعل ما هو أكثر من ذلك ؛ حيث تنتج مخرجات طبقية بالوثائق مرتبة تبعاً لاحتمال اتصالها بموضوع الطلب. وينبغي تقييم نظم الترتيب الطبقي بطريقة تختلف عن تقييم النظم البسيطة ، لأن الحاجة تدعو إلى بيان مدى نجاح عملية الترتيب الطبقي . وتستخدم مقاييس مختلفة لتقييم النظم التي ترتب المخرجات طبقياً، منها « الاستدعاء الطبقي » و « التحقيق المسجل » و « الاستدعاء المعياري » و « التحقيق المعياري » . وهذه المقاييس التي أدخلها سالتون (Salton (1971 تقارن في الأساس الترتيب الطبقي الذي يحققه النظام بالترتيب الطبقي المثالي.

وقد ناقش كين (Keen (1966, 1971 وروبرتسون (Robertson (1969 مقاييس التقييم مناقشة مستفيضة ، وكذلك طرق حساب متوسطات النتائج ، وطرق تقديم

نتائج اختبارات الاسترجاع . ونسب الاستدعاء والتحقيق هي المقاييس الأكثر استخداماً في تقييم نظم استرجاع المعلومات . وقد بسطها كل من كلفردون (Cleverdon 1962) وكلفردون وملز وكين (Cleverdon, Mills and Keen 1966) في مشروع كرانفيلد والأزلب . إلا أن كثيراً من المؤلفين يسوقون بعض أسباب احتمال النظر إلى المقاييس الأخرى باعتبارها أكثر دقة أو أوضح دلالة . ويقدم روبرتسون تحليلاً مفيداً لمزايا وعيوب مختلف المقاييس التي اقترحت أو استخدمت .

ومن بين المعايير المفيدة الأخرى في تقييم نظم استرجاع المعلومات « رقم العمومية » الذي يربط عدد الوثائق المتصلة بموضوع طلب معين بإجمالي عدد الوثائق التي تضمها المجموعة . وبوجه عام ، فإنه كلما ارتفع رقم العمومية (أي كلما زادت كثافة الوثائق المتصلة بالموضوع في مرصد البيانات) ازداد البحث ميلاً إلى السهولة .

وهناك معايير أخرى للكفاءة يمكن استخدامها في تقييم نظم استرجاع المعلومات ، وتشمل « التغطية » و « الجودة » . والتغطية في الواقع امتداد للاستدعاء ، ويتم التعبير عنها بناء على مدى تغطية الإنتاج الفكري لموضوع معين في مرصد بيانات معين . فإذا افترضنا أن باحثاً علمياً يريد العثور على جميع الإشارات الوراقية الممكنة حول الليزر في جراحات العين . فيمكن للكشاف الطبي *Index Medicus* المطبوع أن يكون أحد المصادر البارزة ، وربما كان الأفضل هو خدمة المدلاين الإلكترونية التي تقدمها المكتبة القومية للطب . ونفترض أيضاً أن البحث في مرصد بيانات المدلاين استرجع كل ما هو متصل بالموضوع ، أي حقق نسبة استدعاء ١٠٠ بالمئة ، فإنه حتى إذا كان البحث مكتملاً ، فإن هناك من وجهة نظر مرصد البيانات ، المستفيد الذي يريد بحثاً شاملاً فعلاً ، والذي يرغب أيضاً في معرفة التغطية الفعلية لمرصد البيانات ، أي نسبة إجمالي الإنتاج الفكري حول استخدام الليزر في جراحات العين التي تحظى بالتغطية في مرصد البيانات . فمن الممكن للبحث في مرصد بيانات معين أن يسفر عن ١٠٠ بالمئة استدعاء في الوقت الذي يحقق فيه تغطية إجمالية منخفضة للإنتاج الفكري .

والتغطية ، كالاستدعاء والتحقيق ، يمكن التعبير عنها كنسبة مئوية . فإذا قيّمت نتائج بحث أجري في الـ *Chemical Abstracts* ، فإنه من الممكن أن نقدر بشيء من الصعوبة ، أن نسبة الاستدعاء ٧٥ بالمئة مثلاً ، كما يمكن أن تقدر أيضاً ، وبصعوبة أكثر ، أن تغطية الـ *Chemical Abstracts* للمجال الموضوعي للبحث ٤٠ بالمئة .

وبالتغطية المقدرة بأربعين بالمئة والاستدعاء ٧٥ بالمئة يكون التقدير الكلي لمدى شمول البحث ٣٠ بالمئة ($٠,٤٠ \times ٠,٧٥ \times ١٠٠$ بالمئة) .

ونسبة الجودة ، أي نسبة المواد المتصلة بالموضوع المسترجعة في بحث ما ، والجديدة بالنسبة للمستفيد (أي تعرف عليها لأول مرة عن طريق البحث) ملائمة بوجه خاص في تقييم عمليات بحث الإنتاج الفكري التي تتم لأغراض الإحاطة الجارية ، نظراً لأن خدمة الإحاطة الجارية الجيدة تعرّف المستفيدين بالوثائق قبل أن يتعرفوا عليها بوسيلة أخرى .

وعندما يتم ربط معايير التكلفة بمعايير الجودة تنشأ معايير فعالية التكلفة . وتشمل معايير فعالية التكلفة المحتملة ، والتي يمكن أن تطبق على نظم استرجاع المعلومات تكلفة الوحدة لكل مادة مسترجعة (وثيقة أو إشارة ورقية) متصلة بالموضوع ، وتكلفة الوحدة لكل مادة جديدة مسترجعة متصلة بالموضوع . ومن الممكن قياس التكلفة مباشرة بوحدات مالية أو بالوقت والجهد المستنفذين .

ودقة البيانات معيار آخر للتقييم . ويحل هذا المعيار محل الاستدعاء والتحقيق في تقييم نظم الاسترجاع المصممة للإجابة عن الأسئلة ذات الإجابات الحقائقية التي لا لبس فيها . فالإجابة عن سؤال مثل : ما هي نقطة انصهار ... ؟ إما أن تقدم كاملة وصحيحة أو لا تقدم . وعلى ذلك ، فإن خدمات الرد على الاستفسارات ينبغي أن تقيم بناء على اكتمال ودقة ما تقدمه من بيانات .

إجراء التقييم :

من الضروري من وقت لآخر تقييم نظام أو مرفق استرجاع المعلومات لمعرفة المزيد حول مظاهر القوة ومواطن الضعف ، سواء بالنسبة للمرفق في حد ذاته أو بالنسبة للقائمين على إنشائه وإدارته . وفيما يلي الخطوات الأساسية التي ينطوي عليها مثل هذا التقييم :

- ١ - تحديد مجال التقييم .
- ٢ - وضع برنامج التقييم .
- ٣ - تنفيذ التقييم .
- ٤ - تحليل النتائج وتفسيرها .
- ٥ - تعديل النظام أو الخدمة بناء على نتائج التقييم .

وتتطلب الخطوة الأولى ، تحديد المجال ، إعداد مجموعة دقيقة محكمة من الأسئلة التي ينبغي توجيه التقييم للإجابة عنها (الشكل رقم ٤٢) . وعلى الرغم من تقسيم الأسئلة ، في الشكل ، تبعاً لمكونات نظام الاسترجاع (التكشيف والبحث على سبيل المثال) فإن هناك فئتين عامتين فعلاً من الأسئلة ، تبعاً لما إذا كانت جهود الاستكشاف موجهة نحو عوامل النظام أو العوامل البشرية . وحينما تكون عوامل النظام هي موضوع الاستكشاف فإنه يتم تجميع عناصر بعينها كاستراتيجيات البحث وتعيين المصطلحات الكشفية ، ودراسة هذه العناصر ، ولكن باعتبارها مكونات النظام التي تؤثر في كفاءة الاسترجاع لا أكثر . إلا أنه من الممكن للمستئول عن التقييم أن يرغب في دراسة فعالية عمليات البحث أو قرارات التكشيف ، وربط هذه بوحدة أو أكثر من الخصائص ، كالتمرس في المهمة ، أو المعرفة الموضوعية لمن قاموا بتنفيذ هذه المهام . وعلى ذلك ، فإن تحديد المجال إنما هو في الواقع بيان بما يمكن التعرف عليه تحديداً عن طريق الدراسة .

ومن المهم عند مناقشة مجال التقييم بيان بادئ ذي بدء ما إذا كان التقييم قد قصد به أن يكون وصفيّاً خالصاً ، أو ما إذا كان من الممكن للناتج أن تكون تحليلية وتشخيصية أيضاً ، أي التمييز بين التقييم الكلي macro والتقييم التحليلي micro ، كما فصل كنج وبرايينت (1971) King and Bryant القول . فمن الممكن على سبيل المثال الحصول على مؤشرات الاستدعاء والنسب المئوية للتحقيق ، والخاصة بمجموعة معينة من عمليات البحث التي أجريت في فترة زمنية معينة (تقييم كلي) . وفي مقابل ذلك يمكن دراسة بعض عمليات البحث وما تسترجعه من وثائق للتحقق من الأسباب المحددة لاسترجاع المواد غير المتصلة بالموضوع وعدم استرجاع المتصلة بالموضوع (تقييم تحليلي) . وربما يتبين أنه قد تم استرجاع عدة وثائق غير متصلة بالموضوع استجابة لطلب معين لأنه حدث خطأ في تكشيفها . وحينئذ يمكن الاستفادة من هذه المعلومات لتبرير مراجعة بعض سياسات التكشيف أو تنفيذ بعض أساليب التدريب الجاد للمكشفين . أما إذا كان مسئولو البحث أنفسهم هم بؤرة الاهتمام في الدراسة ، فإنها يمكن أن تبين ما إذا كان المبتدئون من مسئولي البحث يواجهون صعوبات مع النظام بوجه عام ، أو مع نوعيات بعينها من عمليات البحث ، حيث يمكن الاستفادة من هذه المعلومات لإنتاج مواد تدريبية وأدوات أكثر فعالية للتعامل مع الخط المباشر .

الكفاءة بوجه عام :

١ - ما مستوى كفاءة النظام بوجه عام بالنسبة لمتطلبات المستفيدين ؟ وهل هناك اختلافات جوهرية بالنسبة لمختلف أنواع الطلبات وفي مختلف المجالات الموضوعية العريضة ؟

التغطية والتجهيز :

١ - ما مدى سلامة السياسات الحالية الخاصة بتغطية التكشيف ؟
٢ - هل يؤثر الفاصل الزمني بين تلقي الدوريات وتجهيزها بقسم التكشيف ، في الكفاءة بشكل ملحوظ ؟
التكشيف :

١ - هل هناك تفاوت ملحوظ في مستوى كفاءة المكشفين ؟
٢ - إذا كان الأمر كذلك ، فإلى أي مدى يرتبط ذلك بالخبرة ؟
٣ - إلى أي مدى يرتبط بالتألف مع المجال الموضوعي ؟
٤ - إلى أي مدى يرتبط بدرجة المراجعة ؟
٥ - هل يتحقق المكشفون من الموضوعات الدقيقة التي تهم مختلف فئات المستفيدين ؟
٦ - ما أثر السياسات الحالية المتعلقة بالشمول في التكشيف ، على نتائج البحث ؟
لغة التكشيف :

١ - هل المصطلحات مخصصة بما فيه الكفاية ؟
٢ - هل يؤثر تفاوت تخصيص المصطلحات في مختلف المجالات ، بشكل واضح في الكفاءة ؟
٣ - هل هناك ما يدل على الحاجة إلى مزيد من أدوات التحقيق ، كالوزن أو مؤشرات الدور أو رؤوس الموضوعات المركبة ؟
٤ - هل مستوى الترابط بين المصطلحات في المكنز ملائم ؟
٥ - هل لغة الوصول إلى المداخل الكشفية ملائمة ؟
البحث :

١ - ما هي متطلبات المستفيدين بشأن الاستدعاء والتحقيق ؟
٢ - هل يمكن صياغة استراتيجيات بحث تلبي الحاجة إلى الاستدعاء المرتفع أو التحقيق المرتفع ؟
٣ - ما مدى فعالية غرلة المخرجات من جانب مسئول البحث ؟ ما أثر الغرلة على نسب الاستدعاء والتحقيق ؟

٤ - ما هي أفضل سبل التفاعل بين المستفيد والنظام ؟
أ - تكثيف الاتصال في مرحلة الطلب ؟
ب - تكثيف الاتصال في مرحلة صياغة استراتيجية البحث ؟
ج - تواجد المستفيد أثناء إجراء البحث ؟
٥ - ما أثر مختلف طرق التفاعل هذه على زمن الاستجابة ؟
٦ - ما مدى فعالية البحث بواسطة المستفيد النهائي / الوسيط ؟
أ - في ظل طريقة الأوامر ؟
ب - في نظم قوائم الاختيار ؟
٧ - ما أثر مختلف الخصائص المعروفة لمسئولي البحث على كفاءة الاسترجاع ؟

وتتطوي الخطوة الثانية في التقييم ، وهي التخطيط ، على وضع خطة تنفيذية تكفل تجميع البيانات اللازمة للإجابة عن الأسئلة التي طرحت في تحديد مجال التقييم. وعلى مسئول التخطيط للدراسة التحقق من خصائص البيانات اللازمة للإجابة عن كل سؤال، وتحديد الإجراءات التي يمكن اتباعها لتجميع البيانات بأكثر السبل كفاءة وسرعة. وعلى مسئول التقييم أن يحدد لكل سؤال ما إذا كان من الممكن الإجابة عنه ، ببساطة ، بتجميع البيانات من النظام كما هو عليه الآن ، أو ما إذا كان من الضروري إدخال بعض التغييرات على الأداء العادي للنظام لتجميع البيانات اللازمة . وللإجابة عن سؤال مثل « ما هو زمن الاستجابة الحالي للنظام ، معبراً عنه بالمدد والمتوسطات والأنماط ؟ » فإنه ينبغي فقط تجميع البيانات من مسار الأداء العادي للنظام القائم . ولكن للإجابة عن سؤال مثل « ما أثر التكشيف على المستوى المرتفع للشمول ؟ » فإنه يتعين ملاحظة وتوثيق أداء النظام قبل التغير المتعمد في مدى الشمول في التكشيف وبعد هذا التغير. وهكذا يمكن إعداد دراسة وصفية تحليلية لأداء النظام القائم ، أو ربما يتطلب الأمر إجراء تجربة منضبطة حقيقية.

ويوضح الشكل رقم (٤٣) الأسلوب المؤلف في تصميم التجارب . ويتم إجراء التجربة بالتوزيع العشوائي للعناصر موضوع الدراسة على مجموعتين ؛ المجموعة التجريبية والمجموعة الضابطة . ويتم قياس متغيرين ؛ المتغير المستقل أي المعالجة ، ومتغير تابع واحد أو أكثر، وهو خصائص الموضوعات التي يعتقد أنها سوف تتغير بعد المعالجة. وتعرض كل من المجموعة التجريبية والمجموعة الضابطة لاختبار مسبق لقياس المتغيرات التابعة. وتعقب الاختبار المسبق المعالجة (المتغير المستقل) التي تطبق على المجموعة التجريبية فقط. وأخيراً يأتي الاختبار اللاحق ، حيث تقاس المتغيرات التابعة لكل من المجموعتين ثانياً . وإذا كشف قياس الاختبار اللاحق عن اختلاف جوهري بين المتغيرات التابعة للمجموعتين الضابطة والتجريبية ، لم يكن موجوداً في الاختبار المسبق، حينئذ يمكن القول بأن المتغير المستقل كان علامة حاسماً في أداء المجموعة التجريبية.

وبالإضافة إلى التصميم التجريبي الحقيقي الذي عرضنا له توأ ، هناك أيضاً التصميم الارتجاعي *ex post facto* شبه التجريبي . ويتم إجراء هذا النوع من الدراسة عندما يتعذر الحصول على مجموعتين متقابلتين من الموضوعات التي توزع عشوائياً،

كما هو الحال غالباً في مواقف الحياة الفعلية (Campbell and Stanley , 1963, p. 47) .
 ووجه الاختلاف المهم بين هذا النوع من الدراسة والتصميم التجريبي الفعلي ، أن
 مسئول التقييم في الأخير يتحكم في جميع المتغيرات التابعة ، بينما في الأول تكون بعض
 المتغيرات التابعة أو كلها بمنأى عن سيطرة مسئول التقييم^(١) . وهكذا تضعف القوة
 التنبؤية ، أي البرهنة على أن للمتغير المستقل أثراً على أداء الموضوع ، بشكل واضح في
 التجارب الارتجاعية .

| | قيمة المتغير التابع قبل المعالجة | قيمة المتغير التابع بعد المعالجة |
|--------------------|----------------------------------|----------------------------------|
| المجموعة الضابطة | | |
| المجموعة التجريبية | | |

الشكل رقم (٤٣) : تصميم مبسط لتجربة .

ويشتمل الشكل رقم (٤٤) على أمثلة للمتغيرات التي أمكن التحقق منها في عمليات
 تقييم نظم استرجاع المعلومات . وقد سبق أن تناولنا فعلاً متغيرات النظام ومتغيرات
 المخرجات . فضلاً عن ذلك ، هناك الكثير من متغيرات العمليات ومتغيرات الخصائص
 البشرية التي يمكن دراستها . وتتكون متغيرات العمليات عادة من العناصر المكونة
 لدورة عمليات الاسترجاع ، كعدد الأوامر المستقلة المستخدمة في الدورة . وعادة ما
 تتجمع هذه البيانات في سجل للوقائع ، يلتقط بشكل إلكتروني جميع التفاعلات التي
 حدثت بين المستخدم والنظام . ومن الممكن إذن تحليل هذا السجل وتبويبه من جانب

(١) ويمثل هذا في الواقع الحالة المثالية التي يمكن فيها التحقق من جميع المتغيرات التابعة والتحكم فيها . إلا أنه في الواقع
 الأملي ، غالباً ما يكون هناك ، في كل من التجارب الحقيقية وأشياء التجارب ، متغيرات تابعة أخرى ، مجهولة للباحث ،
 ومن ثم غير خاضعة للسيطرة . وتسمى هذه المتغيرات غالباً بالمتغيرات « المربكة » أو « الاعتراضية » . (*)

(*) تسمى هذه في مناهج البحث بالعوامل الملوثة . (المترجم) .

المسئول عن إجراء الدراسة. وفي دراسة الخصائص البشرية يحدد المسئول عن الدراسة بعض الجوانب كالخبرة العملية أو المعرفة الموضوعية التي يعتقد أنها ترتبط ارتباطاً جوهرياً بكفاءة نظام الاسترجاع. وعادة ما يتم تجميع هذه البيانات بالاستبانة.

متغيرات النظام :

شمول الكشف

مستوى الربط في لغة الاسترجاع

مدى التخصيص في البحث

متغيرات الخصائص البشرية :

الخبرة (في البحث أو الكشف)

المعرفة الموضوعية (لجال نظام الاسترجاع)

متغيرات المخرجات :

الاستدعاء

التحقيق

الجدة

حجم المجموعة المسترجعة

الأهمية المعروفة للاسترجاع

متغيرات العمليات :

متغيرات التعلم

زمن التدريب

معدل الأخطاء أثناء التدريب

متغيرات التفاعلية

عدد ونوع الأوامر المستخدمة

الشكل رقم (٤٤) : بعض متغيرات تجربة.

وقد أمكن دراسة كثير من متغيرات نظم استرجاع المعلومات . فالباحث يحدد بوجه عام متغيراً بعينه يمكن أن يكون له أثر على كفاءة الاسترجاع ، وليكن إحدى خطط وزن المصطلحات. وتستخدم مجموعة واحدة (ويفضل أكثر من واحدة) لاختبار استرجاع المعلومات لإجراء التجربة. ويتم تحديد مدى فعالية عمليات البحث الاختبارية ، مقياساً بالاستدعاء والتحقيق، للمجموعات قبل استعمال أداة وزن المصطلحات وبعد استعمال

هذه الأدوات. وهذا التصميم شبه تجريبي ، نظراً لأن المجموعات ليست متناظرة فعلاً. وعلى ذلك، فإن أي إدعاء حول أثر خطة وزن المصطلحات ، ينبغي أن يكون مصحوباً بالاعتراف بالمتغيرات الأخرى التي يمكن أن يكون لها تأثيرها . وهناك بالفعل كثير من المتغيرات، ومن بينها حجم مرصد البيانات ، والمجال الموضوعي وسياسة التكشيف.

ومن الممكن استخدام التصميم التجريبي الحقيقي للتعرف على أثر أحد البرامج التدريبية على كفاءة الاسترجاع الخاصة بالمسئولية عن البحث، حيث يمكن توزيع مسئولي البحث عشوائياً على مجموعتين :

(١) المجموعة التجريبية التي يطبق عليها برنامج التدريب .

(٢) المجموعة الضابطة التي لا تتلقى تدريباً .

وحينئذ يمكن تجميع البيانات الخاصة بفعالية ما تقوم بإجرائه كل من المجموعتين من عمليات البحث، مقيسة بالاستدعاء والتحقيق. فإذا كانت نتائج عمليات البحث التي أجرتها المجموعة التجريبية أفضل بشكل ملحوظ من نتائج تلك التي أجرتها المجموعة الضابطة ، فإنه يمكن القول بأن لبرنامج التدريب أثراً إيجابياً .

وتوضح هذه الخطوات استخدام التصميم التجريبي البسيط جداً والتصميم شبه التجريبي . وللحصول على معالجة أكثر اكتمالاً للطرق التجريبية نحيل القارئ إلى (1963) Campbell and Stanley ، وللحصول على مناقشات تتصل بالمكتبات وعلم المعلومات على وجه التحديد ، نحيله إلى جولدور (1972) Goldhor وبوشا وهارتر (1980) Busha and Harter وباول (1985) Powell .

والخطوة الثالثة ، وهي تنفيذ التقييم ، هي المرحلة التي يتم فيها تجميع البيانات. ويمكن لهذه المرحلة أن تستغرق أطول وقت ممكن. كما أنها يمكن أن تكون المرحلة التي تقل فيها المشاركة المباشرة للمسئول عن التقييم إلى أقصى حد ممكن ، وربما تكون أيضاً المرحلة التي لا تخضع لسيطرته المباشرة إلا في أضيق الحدود.

ويمكن للخطوة الرابعة الخاصة بتحليل النتائج وتفسيرها أن تبدأ قبل الانتهاء من مرحلة التنفيذ. وعلى مسئول التقييم أن يضمن التدفق المستمر للبيانات من بداية مرحلة التنفيذ، حتى يمكن اختزالها في شكل مناسب للتحليل والتفسير. وفي مرحلة

التحليل والتفسير، في مشروع التقييم يقوم المسئول عن التقييم بمعالجة البيانات بطريقة تتيح القدرة على استثمارها في الإجابة عن الأسئلة التي طرحت في خطة العمل.

وفي الدراسة التجريبية ، يتركز الاهتمام في مرحلة التحليل بالمشروع على تفسير أوجه الاختلاف في نتائج الأداء ، كنسب الاستدعاء والتحقيق مثلاً ، في ظل وجود واحدة أو أكثر من الخصائص ، كخبرة مسئول البحث مثلاً ، أو إدخال بعض أدوات لغات الاسترجاع كوزن المصطلحات. وفي الدراسة الوصفية التحليلية لمكونات أحد نظم الاسترجاع العاملة يتم أيضاً الحصول على نتائج الأداء ، وبلي ذلك تحليل أخطاء الاستدعاء والتحقيق. ويستلزم تحليل الأخطاء النظر في كل وثيقة مرتبطة بالموقف، وفي تسجيلات الكشف الخاصة بالوثائق ، وفي الطلبات التي أدت إلى إجراء عمليات البحث، وفي استراتيجيات البحث ، وفي لغة النظام ، وفي قرارات الصلاحية التي يتخذها المستفيدون. وبالنظر في كل هذه العناصر يمكن أن يتقرر أي مكونات النظام كان مسئولاً، إلى حد بعيد، عما حدث من فشل . كذلك يمكن لمسئول التقييم أن يتخذ من نسب الاستدعاء والتحقيق ، أو من مقاييس كفاءة البحث الأخرى ، مؤشرات للظروف التي يعمل النظام في ظلها بصورة جيدة أو بصورة سيئة ؛ فمن الممكن على سبيل المثال تقسيم عمليات البحث إلى مجموعات وفقاً لفئات موضوعية عريضة، والحصول على متوسطات أرقام الأداء الخاصة بكل مجموعة ، وبذلك يمكن التعرف على المجالات الموضوعية التي تحققت فيها أرقام أداء منخفضة بشكل غير عادي.

وعن طريق استعمال أرقام الأداء ، مرتبطة بالاختبارات الإحصائية للدلالة أو بتحليلات الفشل في عمليات بحث معينة ، يمكن لمسئول التقييم معرفة الكثير عن خصائص النظام والمستفيدين منه ، أي مواطن الضعف وأوجه القصور بالإضافة إلى مظاهر القوة. ويمكن لاستعمال كل من أرقام الأداء وتحليلات الفشل أن تجيب عن معظم الأسئلة التي طرحت في خطة العمل الخاصة بالتقييم.

والعنصر الأخير في مرحلة التحليل والتفسير هو الخاص بتقديم مسئول التقييم للنتائج، متضمنة التوصيات الخاصة بما يمكن اتخاذه للارتفاع بمستوى كفاءة النظام. وتأتي بعد هذه المهمة الخطوة الخامسة النهائية في برنامج التقييم ، والتي يتم فيها تنفيذ بعض التوصيات أو كلها (وبذلك تطبق نتائج التقييم لتطوير النظام والمستفيدين منه).

كلمة أخيرة عن التقييم : من المهم قبل إجراء التقييم الكامل، تجربة جميع الإجراءات المقترحة في عينة صغيرة من الواقعات ، وذلك للتأكد من صلاحية الإجراءات للتطبيق فعلاً ، والاطمئنان إلى قدرتها على تجميع البيانات اللازمة لإتمام الدراسة.

الحصول على أرقام الأداء :

ينبغي ألا ننسى أن هناك ثلاثة مستويات يمكن بناء عليها تقييم نظم استرجاع المعلومات ، وهي الفعالية ، رفاعلية التكلفة ، وعائد التكلفة (الشكل رقم ٢٨) . وللتعرف على مدى الوفاء ببعض معايير الفعالية ، وهي على وجه التحديد زمن الاستجابة ، ومقدار الجهد الذي يبذله المستفيد ، والتكلفة المالية ، الوفاء بهذه المعايير من جانب أحد نظم استرجاع المعلومات ، فما على المرء إلا أن يلاحظ الأداء العادي للنظام . أما المعايير النوعية فأكثر تعقداً ، ولا غنى عن اتخاذ خطوات بعينها لتجميع بيانات الأداء.

ويمكن أن يكون محور التركيز في التخطيط لتقييم أحد نظم استرجاع المعلومات أو إحدى خدمات الاسترجاع ، قياس الاستدعاء والتحقيق في عينة ممثلة من عمليات البحث التي يتم إجراؤها. ونسب الاستدعاء والتحقيق (أو بعض الطرق الأخرى للتعبير عن النتائج في الجدول ٢×٢) هي أهم مقاييس جودة عمليات البحث المفوض في أي نوع من النظم. وتدل نسب الاستدعاء والتحقيق مجتمعين على مدى نجاح النظام في غربلة مرصد البيانات لاسترجاع المواد المتصلة بالموضوع وتجنب ما عداها.

وربما كان من الممكن توضيح سبل الحصول على أرقام الأداء هذه ، على أحسن وجه ، بمثال : نفترض أننا بصدد تقييم كفاءة نظام استرجاع معين ، وأن هذا التقييم يتم اعتماداً على عينة عشوائية ، مكونة مثلاً من مئة عملية بحث أجريت خلال فترة زمنية معينة ، وكان النظام يعمل وفق طريقة البحث المفوض ، حيث تقدم طلبات البحث لمركز المعلومات، وتقدم النتائج للمستفيدين في شكل قوائم بالإشارات الوراقية المسترجعة.

دعنا نتتبع إحدى عمليات البحث هذه للتعرف على البيانات التي ينبغي تجميعها، والإجراءات التي يتعين اتباعها لتقييم كفاءة الخدمة بالنسبة لعمليات البحث هذه. وأول ما تدعو الحاجة إليه تسجيل تحريري لطلب المستفيد للمعلومات. ويفضل تسجيل الطلب بلغة المستفيد وأسلوبه ، على نموذج معد خصيصاً لهذا الغرض. وتختلف النماذج في محتواها وطريقة إخراجها، إلا أنها دائماً ما تطلب من المستفيد بيان حاجته إلى المعلومات بأسلوبه وكلماته، وأن يقدم إن أمكن كلمات مفتاحية أو عبارات للدلالة على

المجالات الموضوعية، ويسجل أية مقالات متصلة بالموضوع يعرفها فعلاً ، وأن يعطي أية معلومات أخرى يمكن أن تفيد في تحديد عملية البحث، كاللغة مثلاً ، أو تاريخ النشر، أو نوعية الوثائق. ويشتمل الشكل رقم (١٤٥) على نموذج طلب بحث استخدم في دراسة للتقييم ، أجريت بجامعة متشجان (Janes, 1991). وقد أعدت استراتيجية بحث للطلب الافتراضي بواسطة أحد العاملين بمرفق المعلومات (الشكل رقم ٤٥ ب)، واسترجع البحث ٢٥ إشارة وراقية من مرصد البيانات . وتدعو الحاجة إلى وجود نسخة من استراتيجية البحث لأغراض التقييم .

(١)

مدرسة المعلومات ودراسة المكتبات
دراسة الصلاحية
نموذج طلب بحث

نرجو تقديم وصف سردي موجز لموضوعك (استعمل ظهر النموذج إذا دعت الضرورة)
أنا مهتم بتأثير التغيرات الجوية على الحالة المزاجية . وبمزيد من التحديد موضوع الاكتئاب الفصلي أو الاكتئاب الشتوي ، الذي يعتقد أنه يرجع إلى تناقص ضوء الشمس (تغير جوي) خلال الشتاء . وأود الحصول على معلومات عن الدراسات التي استكشفت الارتباط بين الطقس والحالة المزاجية . وربما يرد كثير من هذه الدراسات وليست جميعها تحت موضوع الاكتئاب الفصلي أو الشتوي .

هل تعرف أي مصطلحات كشفية ، أو مصطلحات وردت في لغة النظام ، أو مصطلحات بحث ، يمكن أن تفيد في البحث عن الوثائق في هذا الموضوع ؟ نرجو تسجيلها هنا أو ضع خطأ تحتها في الوصف الوارد أعلاه .

| | |
|----------|-----------|
| المزاج | الفصلي |
| الانفعال | الشتاء |
| الشعور | و الطقس |
| الاكتئاب | ضوء الشمس |

هل تعرف أيًا من المؤلفين أو الوثائق المتصلة بهذا الموضوع ؟ نرجو بيانها هنا :

N. E. Rosenthal
T. A. Wehr
B. Ainswirth

M. E. Thase
C. S. Pittendrigh

نوعيات الوثائق التي تهتمك (ضع دائرة) :

| | | | |
|-----------------|----|----------------|----|
| مقالات الدوريات | لا | بحوث المؤتمرات | لا |
| التقارير | لا | الاطروحات | لا |

نوعيات أخرى (حددها) : فصول الكتب

الأعوام المطلوب تغطيتها : الأعوام الثلاثون الماضية

اللغات التي تهتمك : الإنجليزية فقط

نرجو تسجيل أي معلومات أخرى يمكن أن تراها مفيدة في صياغة استراتيجية البحث ، على الظاهر.

الشكل رقم (٤٥) : ١ مثال لنموذج طلب بحث.

(ب)

SEASONAL(W)AFFECTIVE(W)DISORDER?

يسفر عن

S1

مجموعة الوثائق المسترجعة

SEASONAL(W)VARIATION? and AFFECTIVE(W)DISTURBANCE?

يسفر عن

S2,

الموضوع بالربط بين الجانب الفصلي وجانب المزاج

S2, m S2

مجموعة الوثائق المسترجعة

يسفر عن

S3

الاتحاد المنطقي للمجموعتين الأوليين

S3/eng

المجموعة الناتجة عن الاتحاد المنطقي

S4

المجموعة مقتصرة على الوثائق بالإنجليزية

S4/human

مجموعة الوثائق الناتجة

يسفر عن

S5

تقتصر على الدراسات المهمة بالبشر

تدل الخطوط السميكة على وجود بعض البيانات

ولحساب نسبة التحقيق لعملية البحث ينبغي أن يطلب من المستفيد تحديد أي المواد المسترجعة صالحة وأيها ليست كذلك . ويفضل إصدار أحكام الصلاحية هذه بناء على الوثائق نفسها ، أو على الأقل بناء على بدائل مكتملة للوثائق كالمستخلصات مثلاً . وينبغي تعبئة نموذج تقدير صلاحية لكل وثيقة يتم تقييمها . ويشتمل الشكل رقم (٤٦) على مثال لهذا النموذج (كما استخدمه لانكستر [1968 a] Lancaster في تقييم المدايرز MEDLARS) والذي يلتصق أساساً استجابات حول جودة كل وثيقة (السؤال رقم ١) ومدى صلاحيتها (السؤال رقم ٢) . ومن المهم أن يطلب من المستفيد تسجيل أحكامه على مدرج قيم يحدد مختلف درجات الصلاحية ، نظراً لأن ذلك يعبر عن أحكام المستفيد بدقة لا يكفلها التقسيم الثنائي البسيط صالح / غير صالح . وعلى المستفيد أن يسجل أسباب الحكم على كل وثيقة بالصلاحية أو عدم الصلاحية . وإذا استرجع البحث عدداً كبيراً من الوثائق ، فإنه يكفي لأغراض التقييم أن يطلب من المستفيد الحكم على صلاحية عينة عشوائية فقط من الوثائق المسترجعة .

المكتبة القومية للطب
بسيڊا - مريلاڊ

طلب رقم _____
وثيقة رقم _____

مشروع تقييم المدلرز
نموذج تقييم الوثائق

- ١ - هل كنت من قبل على دراية بوجود هذه المقالة ؟
نعم [] كيف علمت بوجودها ؟
لا []
- ٢ - نرجو تقييم هذه المقالة من حيث علاقتها بالحاجة إلى المعلومات التي أدت إلى تقديم الطلب للمدلرز، وذلك بوضع علامة في الخانة المناسبة .
(أ) مهمة جداً لي من حيث اتصالها بحاجتي إلى المعلومات []
وضح الأسباب رجاء :
(ب) قليلة الأهمية لي من حيث اتصالها بحاجتي إلى المعلومات []
وضح الأسباب رجاء :
(جـ) لا أهمية لها بالنسبة لي من حيث اتصالها بحاجتي إلى المعلومات []
وضح الأسباب رجاء :
هل كنت سعيداً بمعرفة وجودها بسبب حاجة أخرى أو مشروع آخر ؟
نعم [] وضح الأسباب رجاء :
لا []
(د) غير قادر على التقييم بسبب لغة الوثيقة []
هل تنوي اتخاذ أي خطوات لمعرفة محتوى هذه الوثيقة الصادرة بلغة أجنبية ؟
نعم [] حدد الخطوات رجاء :
لا [] بين الأسباب رجاء :

الشكل رقم (٤٦) : نموذج تقييم الصلاحية في المدلرز.

وإذا رجعنا إلى البحث الافتراضي الذي استرجع ٢٥ وثيقة ، وافترضنا أن المستفيد حكم بصلاحية ١٥ وثيقة وعدم صلاحية ١٠ وثائق ، وافترضنا أيضاً أن مرصد البيانات برمته يضم ٥٠٠,٠٠٠ وثيقة ، فإنه من الممكن وضع قيم معينة في الجدول ٢×٢ الخاص بنتائج البحث على النحو التالي :

| صالح | غير صالح | المجموع |
|------------|----------|---------|
| ١٥ | ١٠ | ٢٥ |
| ٤٩٩,٩٧٥ | | |
| ٥٠٠,٠٠٠ | | |
| مسترجع | | |
| غير مسترجع | | |
| المجموع | | |

أما نسبة التحقيق الخاصة بعملية البحث فهي ١٥/٢٥ أي ٦٠ بالمئة .

والمعضلة الأساسية الباقية هي تقدير الاستدعاء . وقد استعمل مصطلح «تقدير» هنا عمداً لأنه لا يمكن تحديد قيمة الاستدعاء بالنسبة لعملية البحث بشكل نهائي جازم، إلا إذا كان المستفيد على استعداد للنظر في جميع التسجيلات التي لم تسترجع والبالغ عددها هنا ٤٩٩,٩٧٥ تسجيلة ، وأن يبين ، وفقاً لنفس المعايير التي سبق أن اتبعها أي هذه التسجيلات صالح أو متصل بالموضوع وأياً ليس كذلك . كما أنه لا يمكن التوصل إلى تقدير للاستدعاء بأخذ عينات عشوائية بطريقة تقليدية من المواد التي لم تسترجع، لأن مجموعة المواد غير الصالحة بالنسبة لطلب معين عادة ما تكون أكبر بكثير من مجموعة المواد الصالحة. ومن ثم فإن الأمر قد يتطلب سحب عينة عشوائية ضخمة إلى حد لا يحتمل، من ٤٩٩,٩٧٥ مادة لم تسترجع في هذا المثال ، لكي يكون أمامنا احتمال العثور على مجرد وثيقة واحدة صالحة في العينة . (إلا أن بلير Blair [1990] طور طريقة لتقدير الاستدعاء باستعمال « أطر سحب العينات » ، نتناولها فيما بعد).

وتنطوي إحدى الطرق المحتملة لتقدير الاستدعاء على استخدام عمليات بحث إضافية يجريها أعضاء آخرون من العاملين بمركز المعلومات. ويتم التعبير عن استدعاء عملية البحث الأصلية التي أجراها العضو أ على النحو التالي :

عدد المواد الصالحة التي عثر عليها أ

عدد المواد الصالحة التي عثر عليها أ + عدد المواد الصالحة التي عثر عليها بمرصد البيانات بواسطة ب ، و ج ، ...

ويعني ذلك ضرورة تقديم أية وثائق إضافية عثر عليها ب ، و ج وغيرهما ولم يعثر عليها في عملية البحث الأصلية ، إلى المستفيد للحكم على صلاحيتها .

ويمكن « للاستدعاء المقارن » أن يكون كافياً بالنسبة لبعض أغراض التقييم ؛ فإذا كان هناك على سبيل المثال ، من يريد التعرف على مدى براعة الكيميائيين في البحث في مرصد البيانات الكيميائية على الخط المباشر لتلبية احتياجاتهم من المعلومات ، مقارنة بالاعتماد على اختصاصي المعلومات في إجراء عمليات البحث ، فإنه من الممكن التعبير عن استدعاء البحث الذي أجراه الكيميائي على النحو التالي :

عدد الوثائق الصالحة التي عثر عليها الكيميائي

عدد الوثائق الصالحة التي عثر عليها الكيميائي + عدد الوثائق
الصالحة الإضافية التي عثر عليها اختصاصي المعلومات

ومن الملاحظ أنه لأغراض التقييم التحليلي التشخيصي ، ينطوي تقدير نسبة الاستدعاء بإجراء المزيد من عمليات البحث في مرصد البيانات نفسه ، على عيب بارز ، وهو محدودية نوعيات أخطاء الاستدعاء التي يمكن أن يكشف عنها . فمن الممكن لهذا الأسلوب أن يكشف عن أخطاء الاستدعاء الراجعة إلى ضعف استراتيجيات البحث ، أو قصور اللغة ، أو الخلل في تفاعل المستفيد مع النظام ، إلا أنه قد لا يكشف أخطاء التكشيف . فإذا كانت هناك وثيقة تهتم بالموضوع ص إلا أن الموضوع ص لم يكشف ، فإنه لا يمكن لمسئول البحث أ أن يسترجع الوثيقة عند البحث عن ص . كما أنه لا يمكن أيضاً لأي مسئول بحث آخر العثور على هذه الوثيقة . ومما لا شك فيه أن هذه الطريقة لا تكفل لنا سوى تقدير الحد الأعلى لاستدعاء عملية البحث الأصلية ، نظراً لأن جميع مسئولو البحث مجتمعين يمكن أن تفلت منهم بعض الوثائق المفيدة . ونظراً لأنه لا يمكن لاثنتين من مسئولو البحث العثور على كل ما هو صالح في مرصد البيانات (ولقد تبين لساراسفك وكاننتور [1988 b] Saracevic and Kantor فعلاً أنه لا يمكن حتى لستة أو أكثر من مسئولو البحث العثور على كل شيء) فإن تقدير الاستدعاء الذي حققه الكيميائي ، لا يمكن أن يكون سوى تقدير للحد الأعلى . وعلى ذلك ، فإنه إذا عثر الكيميائي على عشر مواد صالحة ، وعثر اختصاصي المعلومات على خمس آخر ، فإن استدعاء الأول لا يمكن أن يكون أفضل من ١٥ / ١٠ ، وربما يكون أسوأ .

وهناك أسلوب أفضل من ذلك لتقدير الاستدعاء في النظام العامل فعلاً ، وهو إجراء بحث مواز في مرصد آخر للبيانات أو أكثر ، ويمكن لهذه الأخرى أن تكون كشافات

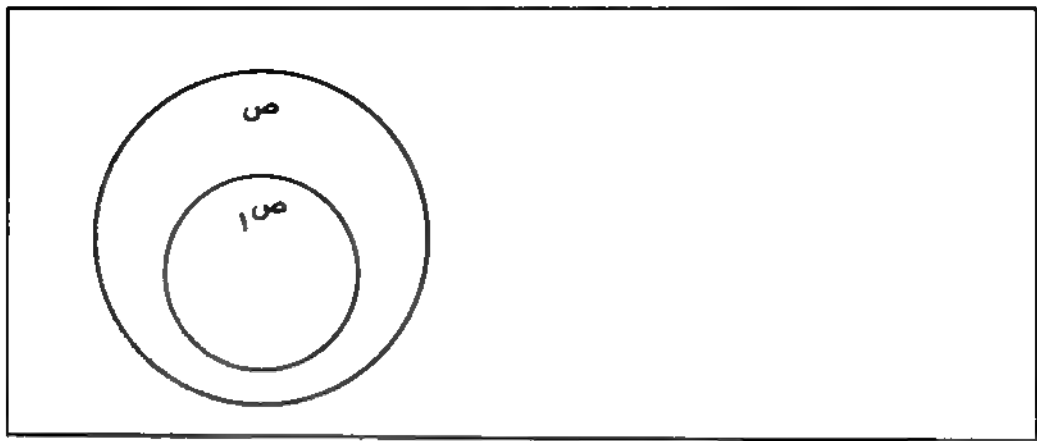
مطبوعة، كما فعل لانكستر (1968 a) Lancaster في تقييم المدلرز : هب أن هناك من يرغب في تقدير نسبة الاستدعاء الخاص ببحث للإنتاج الفكري في مرصد بيانات معين، في موضوع التلوث الحراري للمياه ، فيتم إجراء البحث نفسه في مرصد بيانات أخرى ، حيث يمكن الاعتماد على أي مرصد للبيانات مادام مستقلاً تمام الاستقلال عن المرصد الذي أجري فيه البحث الأصلي . ونفترض أن البحث الذي لم يكن من الضروري أن يكون شاملاً، أسفر عن استرجاع ١٤ وثيقة بدت متصلة بموضوع التلوث الحراري للمياه ، وقدمت هذه الوثائق إلى المستفيد صاحب البحث الأصلي ، حيث يحكم بصلاحيته ١٢ وثيقة. حينئذ تراجع هذه الوثائق الاثنتا عشرة مقابل مرصد بيانات النظام موضوع التقييم، حيث يتبين اشتماله على عشر وثائق منها . ثم تراجع هذه الوثائق العشر مقابل الناتج المطبوع للبحث الأصلي في النظام ، حيث يتم استرجاع سبع وثائق . ومن ثم فإن تقديرنا للاستدعاء هو ١٠ / ٧ ، أي ٧٠ بالمئة.

وهذا الأسلوب أحد أساليب التقدير الاستقرائي، من مجتمع الوثائق الصالحة المعروفة، لمجتمع الوثائق الصالحة غير المعروفة (الشكل رقم (٤٧)). فبالنسبة لأي طلب بعينه يقدم للنظام ، يشتمل مرصد البيانات الذي يمثله المستطيل بأكمله في الشكل ، على مجموعة فرعية من الوثائق ص التي يمكن للمستفيد أن يحكم بصلاحياتها إذا اطلع عليها ، إلا أن الهوية الكاملة للمجموعة ص غير معلومة . ويؤدي أسلوب التقدير الاستقرائي للتحقق من مجموعة فرعية من ص ، وهي ص١ ، حيث تبين أن نسبة استدعاء ص١ كانت ١٠ / ٧ ، فإنه من المعقول إذن أن نفترض أنه إذا كانت ص١ تمثل ص تمثيلاً دقيقاً ، فإن نسبة الاستدعاء في ص (غير المعروفة) يمكن أن تقارب نسبة الاستدعاء في ص١ (المعروفة) . ويقدم هذا الأسلوب تقديراً للاستدعاء في ص وليس الاستدعاء المؤكد، إلا أن هذا هو أفضل الاحتمالات في ظل معظم الظروف. ويمكن للتقدير أن يفضي أيضاً إلى القيم الأخرى للجدول ٢×٢ الخاص بنتائج عملية البحث. ويدل تقدير الاستدعاء في المثال على أن الوثائق الخمس عشرة الصالحة المسترجعة تشكل حوالي ٧٠ بالمئة من مجموع الوثائق الصالحة في مرصد البيانات ، ومن ثم فإن ١٠٠ بالمئة يمكن أن تكون حوالي ٢١ وثيقة. وبذلك تصبح القيم الكاملة لجدول نتائج عملية البحث على النحو التالي :

| صالح | غير صالح | المجموع |
|------------|----------|---------|
| ١٥ | ١٠ | ٢٥ |
| *٦ | *٤٩٩,٩٦٩ | ٤٩٩,٩٧٥ |
| *٢١ | *٤٩٩,٩٧٩ | ٥٠٠,٠٠٠ |
| مسترجع | | |
| غير مسترجع | | |
| المجموع | | |

والقيم المميزة بنجمة تقديرات ، أما القيم الأخرى في الجدول فمؤكدّة.

وهناك طريقة أخرى لتقدير الاستدعاء في النظم العاملة الضخمة ، تعتمد على أساليب غير تقليدية لسحب العينات ؛ فبدلاً من سحب العينات عشوائياً من مجموعة الوثائق بأكملها ، تطورت طرق لسحب العينات العشوائية من « أطر العينات » هذه الأطر مجموعات فرعية من مرصد البيانات يُتوقع لها أن تكون غنية بالوثائق الصالحة (Blair, 1990) ، حيث يتم تقدير الاستدعاء بناء على عينة عشوائية من هذه المجموعات الفرعية.



الشكل رقم (٤٧) : تقدير الاستدعاء باستقراء النتائج من مجتمع معلوم ص١ إلى غير المعلوم ص.

وهناك أساساً طريقتان لتحديد هذه الأطر ؛ الأولى تعتمد على التعديل المنطقي لاستراتيجية البحث الأصلية ؛ فإذا كان الطلب لكل من « أ و ب و ج و د » مثلاً ، فإنه يمكن إجراء عمليات البحث الأكثر عمومية التالية ، والاعتماد على المجموعات الناتجة لسحب العينة :

- أ و ب و ج وفيما عدا د
 أ و ب و د وفيما عدا ج
 أ و ج و د وفيما عدا ب
 ب و ج و د وفيما عدا أ

أما الطريقة الثانية لتحديد هذه الأطر فهي استعمال مترادفات إضافية في استراتيجية البحث الأصلية ، لإجراء بحث أكثر شمولاً ، مما يسفر عن تجمع من الوثائق أكبر ، يمكن أن يكون بعضها صالحاً . ويعترف بلير (Blair (1990 بأن أساليب سحب العينات هذه عادة ما تسفر عن بيان الحد الأقصى الممكن للاستدعاء الخاص بالبحث الأصلي لا عن تقدير للاستدعاء الحقيقي .

واعتماداً على الإجراءات التي عرضنا لها تَوْأ ، يمكن الخروج بأرقام الأداء الخاصة بعينة ممثلة لعمليات البحث التي أجريت في نظام للاسترجاع على الخط المباشر ، يعمل بطريقة التفويض . أما تقييم البحث الذي يتم على الخط المباشر دون تفويض فموقف مختلف . فما زال علينا تقدير الاستدعاء اعتماداً على أي من الأساليب التي سبق أن تناولناها ، أما نسبة التحقيق فمقياس أقل أهمية من المقياس المباشر لتكلفة الوحدة لكل وثيقة صالحة مسترجعة . وربما كان من الممكن وضع تكلفة البحث محسوبة بالدولار أو بوقت المستفيد ، لا نسبة التحقيق ، في موازنة مقابل نسبة الاستدعاء .

وفي حالة استعمال نماذج تقدير الصلاحية من النوعية الواردة في الشكل رقم (٤٦) ، فإنه يمكن حساب نسبة الجودة ونسبة التحقيق لكل عملية بحث . ويمكن التعبير عن نسبة الجودة بإحدى طريقتين :

عدد الوثائق الصالحة الجديدة المسترجعة

عدد الوثائق الصالحة المسترجعة

أو عدد الوثائق الصالحة الجديدة المسترجعة

عدد الوثائق المسترجعة

هذا ، ويتطلب قياس تغطية مرصد البيانات أسلوباً مختلفاً للتقييم ، أسلوب يعتمد على الوراقيات المتخصصة ، والمراجعات العلمية مصدر جيد لهذه الوراقيات . فإذا افترضنا أن هناك من يريد التعرف على مدى شمول تغطية الكشف الطبي *Index Medicus* لموضوع اضطرابات التغذية ، فإنه يمكن العثور على عدة مراجعات علمية حول جوانب معينة من الموضوع في وراقية *Bibliography of Medical Reviews* . وكلما زاد عدد المراجعات العلمية التي يتم العثور عليها ، وكلما اكتملت تغطية هذه المراجعات ، كانت أفضل في تلبية أغراض التقييم . وإذا افترضنا العثور على ثلاث مراجعات حديثة ، وهذه المراجعات مجتمعة تستشهد بمئة وعشرين وثيقة متفردة في مختلف المصادر ، فإنه تتم مراجعة كل إشارة من الإشارات الوراقية مقابل كشف المؤلف في الـ *Index Medicus* لتحديد أي المواد يشتمل عليها وأيها لا يشتمل عليها ، إلى أن يصبح من الممكن معرفة مقدار ما يغطيه الكشف من الإشارات المئة والعشرين . ومن الممكن العثور على أمثلة لتطبيق هذا الأسلوب في مقالتي مارتن (1967) Martyn ومارتن وسليتر (1964) Martyn and Slater .

تفسير النتائج :

يمكن للتقييم أن يتخذ شكل التجربة المنضبطة أو وصف الوضع الراهن لتشغيل النظام أو نشاط المستفيدين . وفي التجربة تكفل الاختلافات ذات الدلالة الإحصائية بين المجموعة التجريبية والمجموعة الضابطة ، للباحث استنتاج ما إذا كان إدخال المتغير المستقل قد أثر ، بالإيجاب أو بالسلب . وعلى الرغم من قوة التجارب الواضحة ، وما لها من جدوى ، وخاصة في قدرتها على بيان الأسباب ، هناك بعض النقاط الجديرة بالتسجيل حول استخدامها في تقييم نظم استرجاع المعلومات :

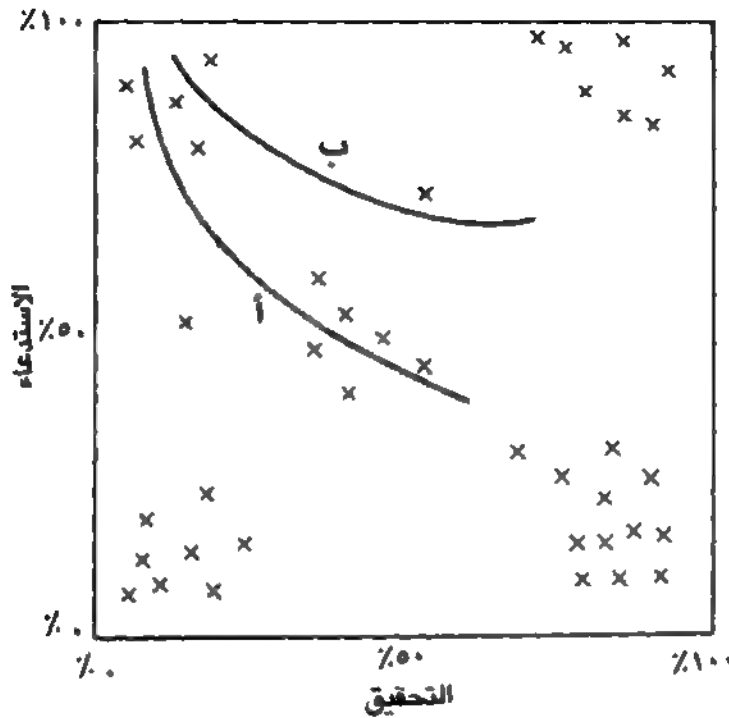
١ - غالباً ما يكون عدد الأفراد موضوع التجربة قليلاً ، نظراً للجهد المكثف المستنفد في تصميم التجربة وتنفيذها . وأحياناً ما يكون عدد من تشملهم الدراسة محدوداً جداً بشكل يحد من الدلالة الإحصائية للاختبارات .

٢ - على الباحث أن يتأكد من أن الظروف الاصطناعية للتجربة لا تؤثر في النتائج ، وأن يأخذ في حسابه أي مؤثرات يمكن أن تغير في النتائج .

٣ - يصعب في غالب الأحيان تجميع مجموعتين (تجريبية وضابطة) متناظرتين فعلاً . ومن ثم ، فإنه يمكن في بعض الأحيان رد مظاهر التفاوت في الأداء إلى الاختلاف بين المجموعتين لا إلى إدخال المتغير المستقل .

أما في الدراسات الوصفية التحليلية ، التي غالباً ما تستخدم في تقييم كفاءة نظم الاسترجاع الضخمة ، فإن الباحث يستخدم مقاييس المخرجات، ثم يحاول بعد ذلك التعرف على كفاءة النظام بتحليل الأخطاء. ومن الممكن الخروج بمتوسط نسب الاستدعاء والتحقيق التي أحرزها النظام بحساب متوسط أرقام أداء عمليات البحث الاختبارية ، للوصول إلى معدلات أداء كلية . وبتعديل استراتيجية البحث في مجموعة من عمليات البحث الاختبارية ، يمكن استخلاص سلسلة من معدلات الأداء، وتوقيع منحني متوسط الأداء ، الذي يشبه إلى حد بعيد المنحنى الوارد في الشكل رقم (٤١) .

ويبين هذا المنحنى متوسط أداء النظام في الوقت الراهن. وعلى مسئول التقييم أن يدرك، بالطبع، أن المتوسطات مضللة، ومن ثم فإن النقاط التي يحرزها الأداء لا ينبغي أن توقع كمنحنى فقط وإنما في شكل بياني للتشتت ، كما هو موضح على نحو مبسط جداً في الشكل رقم (٤٨) . وهنا يمثل المنحنى أ منحني الأداء المتوسط الحالي، وترمز كل × لنقطة أداء بعينها . ويلاحظ أنه لا يقع أي من أرقام الأداء على منحني متوسط الأداء تماماً، حيث نجد في الواقع تشتت نتائج الأداء بشكل ملحوظ. وبعض النتائج جيد جداً (الركن العلوي الأيمن) وبعضها سيء جداً (الركن الأسفل الأيسر) . وبعض النتائج تبدي ارتفاعاً في الاستدعاء وانخفاضاً في التحقيق ، بينما تبدي أخرى ارتفاعاً في التحقيق وانخفاضاً في الاستدعاء ، وكثير منها يقع « في منتصف الطريق » .



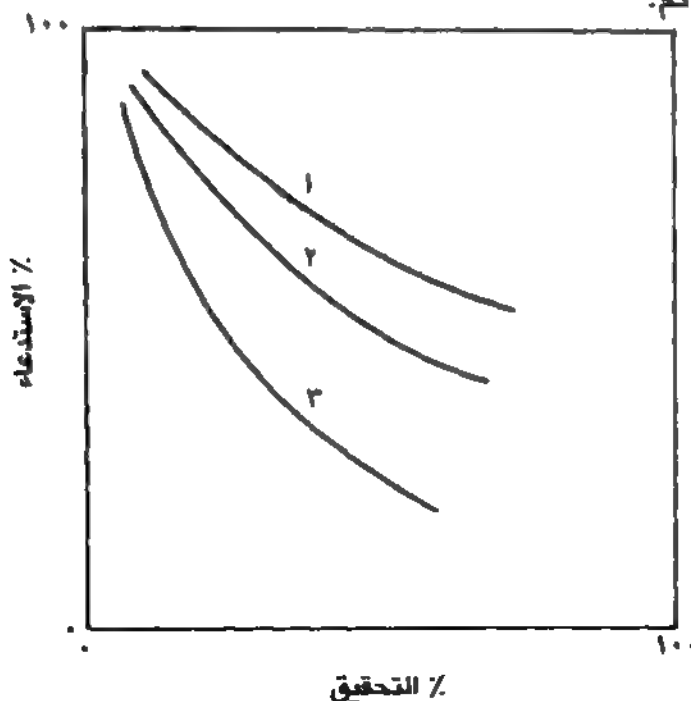
الشكل رقم (٤٨) : رسم بياني لتشتت نتائج عمليات البحث.

وأهم المهام في التقييم هي تمييز مظاهر النجاح من أعراض الفشل ؛ فبتحديد هذه للعوامل التي تجعل البحث جيداً أو سيئاً ، يمكن لمسئول التقييم التحقق من أهم مشكلات النظام واقتراح الحلول الممكنة . وإذا أمكن نتيجة لذلك الارتفاع بمستوى بعض عمليات البحث التي كانت سيئة ، حينئذ يمكن لمتوسط مستوى أداء النظام أن يرتفع ، أي أنه يمكن لمنحنى متوسط الأداء الارتفاع إلى نقطة أعلى ، قريبة من المثالية (وإن كان ذلك لا يمكن بلوغه) في الركن الأيمن العلوي للرسم البياني ، كما هو موضح بالمنحنى ب في الشكل رقم (٤٨) .

وللتعرف على أهم مصادر ضعف النظام يتم تحليل الأخطاء الفردية . وتستعمل أرقام الأداء لمقارنة كفاءة النظام في ظل ظروف متفاوتة أو طرق تشغيل مختلفة . (ولا يمكن استعمال هذه الأرقام لمقارنة كفاءة نظام ما بكفاءة نظام آخر مختلف في وثائقه وفيما يقدم له من طلبات وفيمن يتعامل معه من مستفيدين) . ويمكن تقسيم عمليات البحث الاختبارية بعدة طرق ، بحيث يمكن لأرقام الأداء أن تكشف عما إذا كانت هناك اختلافات جوهرية في الأداء تبعاً لاختلاف ظروف التشغيل . فهل يعمل النظام ، على سبيل المثال ، في المتوسط ، في مجالات موضوعية معينة ، على نحو أفضل مما هو عليه في مجالات أخرى؟ هل يعمل بالنسبة لفئات معينة من المستفيدين أفضل مما يعمل لغيرهم؟ وعلى الرغم من أن أرقام الأداء تدل على حدوث أشياء معينة في النظام ، فإنها لا تبين أسباب ذلك . ويتطلب ذلك تحليلاً فكرياً وتفسيراً .

وبتقسيم نتائج التقييم بعدة طرق ، يمكن الخروج بمجموعة من منحنيات الأداء ، التي يعبر كل منها عن أداء النظام في ظل ظروف معينة للإفادة . ويوضح الشكل رقم (٤٩) مثل هذه المجموعة من المنحنيات . (لاحظ أن الأمر يتطلب إجراء عدد كبير نسبياً من عمليات البحث ، لإجراء نوعية المقارنة التي يمثلها هذا الشكل ، بأي مستوى من مستويات الثقة الإحصائية) . وقد أحرزت المجموعة ١ أفضل نتائج ، بينما سجلت المجموعة ٣ أسوأ نتائج . ويمكن للمنحنيات الواردة في الشكل أن تمثل عمليات البحث التي أجريت في مجالات موضوعية مختلفة ، أو عمليات البحث التي أجريت لنوعيات مختلفة من المستفيدين ، وهكذا . ويمكن لتحليل من هذا النوع أن يبين مواطن الضعف في النظام ، والسبل الكفيلة بالارتفاع بمستوى الأداء . فإذا كان منحنى المجموعة ٣ ، على

سبيل المثال ، يمثل عمليات البحث في مجال موضوعي بعينه ، فإن هذه النتائج السيئة يمكن أن تدل على قصور في لغة النظام في ذلك المجال. فقد كشف التقييم إذن عن الضعف، وبذلك يكفل لمسئول ما (وهو في هذه الحالة منتج مرصد البيانات) اتخاذ الإجراء العلاجي اللازم.



الشكل رقم (٤٩) : نتائج الاستدعاء والتحقيق لثلاث مجموعات من عمليات البحث.

ولنسب الاستدعاء والتحقيق وجه مهم آخر للإفادة منها ؛ فمن الممكن لكل نسبة أن تدل على عدد معين من حالات الإخفاق، ويكفل ذلك لمسئول البحث القدرة على إجراء تحليل للتعرف على أسباب هذا الإخفاق. فإذا نظرنا في عملية بحث افتراضية استرجع فيها النظام ستاً من الوثائق الصالحة العشر المعروفة ، وأفلتت منه أربع ، فإن ذلك يعني أن نسبة الاستدعاء ٦٠ بالمئة . وينظر المستفيد في عينة عشوائية من ٢٥ مقالة مسترجعة ويحكم بصلاحية عشر وبعدد صلاحية خمس عشرة ، حيث يعني ذلك أن نسبة التحقيق ٤٠ بالمئة . ومن ثم فإن التحليل في عملية البحث هذه بالذات ينبغي أن يكون لأربعة أخطاء في الاستدعاء وخمسة عشر خطأ في التحقيق . ولا ننسى أن أخطاء الاستدعاء الأربعة وأخطاء التحقيق الخمسة عشر ، ليست هي فقط الأخطاء التي وقعت في عملية البحث ؛ فهي ليست سوى الأخطاء التي أمكن التعرف عليها لا أكثر ، وطالما هي كذلك فقد قبلت باعتبارها أمثلة لأخطاء الاستدعاء والتحقيق الكاملة لعملية البحث، أي أنها أعراض لمشكلات حدثت في عملية البحث.

والتحليل اللاحق لأخطاء البحث هو أصعب جوانب عملية التقييم، فهو ينطوي بالنسبة لكل خطأ ، على تفحص ما يلي:

- ١ - النص الكامل للوثيقة نفسها.
- ٢ - المصطلحات الكشفية التي خصصت للوثيقة.
- ٣ - بيان الطلب.
- ٤ - استراتيجية البحث التي اتبعت.
- ٥ - نموذج التقييم الذي عبأه المستفيد (ومن العناصر ذات الأهمية الخاصة في تحليل أخطاء التحقيق أسباب الحكم على المقالة بأنها لا قيمة لها).

وبناء على هذه السجلات يتم اتخاذ القرار بصدد الأسباب الرئيسة لوقوع خطأ معين. ومن الممكن رد جميع الأخطاء تقريباً إلى بعض جوانب التكشيف، أو البحث، أو لغة التكشيف، أو قطاع التفاعل بين المستفيد والنظام. وفي الدراسة محكمة التصميم، عادة ما يكون تحليل أخطاء التحقيق على الأقل، مسئولية مشتركة بين المستفيد ومسئول التقييم، لأن بيان المستفيد لأسباب الحكم بعدم صلاحية وثيقة معينة غالباً ما يكون دليلاً جيداً لمعرفة موطن الزلل في النظام. لنفترض على سبيل المثال، أن المستفيد قد أشار إلى أن الوثيقة غير صالحة لأنها تتناول المولدات الإلكترونية للضوضاء ، بينما كان يريد المولدات الميكانيكية للضوضاء . ويكفل مثل هذا البيان لمسئول التقييم أن يعرف على وجه التحديد لماذا فشلت الوثيقة المسترجعة في تلبية حاجة المستفيد إلى المعلومات. وهنا ينبغي فحص السجلات الضرورية لمعرفة ما إذا كانت عملية البحث قد أجريت بشكل عريض جداً، أو ما إذا كانت لغة التكشيف تفتقر إلى التخصيص المناسب للطلب ، أو ما إذا كان قد حدث خطأ في تكشيف الوثيقة ، أو ما إذا كان بيان الطلب يفتقر إلى التحديد الدقيق (كأن يكون قد عجز عن التعبير عن اقتصار الاهتمام على المولدات الميكانيكية للضوضاء دون غيرها).

وحيثما أمكن ، فإنه عادة ما يتم التركيز على أهم عامل وراء كل خطأ على حدة، إلا أنه لا يتسنى في بعض الأحيان التحقق من سبب واحد دون غيره ، نظراً لاشتراك وظيفتين من وظائف النظام على قدم المساواة. وبالنسبة لبعض أخطاء الاستدعاء ، يمكن القول بأنه كان من الممكن للوثيقة أن تسترجع لو أن المكشف استعمل المصطلح الإضافي

١١. إلا أنه لا يقل عن ذلك أهمية القول بأنه كان من الممكن للوثيقة أن تسترجع لو أن مسئول البحث غير من الاستراتيجية المخصصة التي اتبعها (أ و ب و ج) إلى الاستراتيجية المعقولة (أ و ب و ج). وينبغي في مثل هذه الحالات رد الخطأ إلى كل من الكشف والبحث معاً، أو إلى أي من عناصر النظام الأخرى التي كانت مسئولة معاً في الوقت نفسه.

وفي تقييم المدلرز (Lancaster, 1968 a)، حيث أجريت ٢٠٢ عملية بحث، تم تحليل ٧٩٧ خطأ في الاستدعاء و ٣٠٣٨ خطأ في التحقيق. وكانت الأخطاء ترجع إلى المكونات الأساسية للنظام، كما هو موضح في الشكل رقم (٥٠). وتختلف نسب الأخطاء وأنواعها المحددة في كل فئة من نظام إلى آخر. إلا أنه يمكن القول إن الأنواع الرئيسية لأخطاء الاسترجاع واحدة بالنسبة لمعظم نظم الاسترجاع العاملة فعلاً.

| أخطاء التحقيق | أخطاء الاستدعاء | |
|---------------|-----------------|-----------------------------------|
| ١٠٩٤ (٣٦,٠٪) | ٨١ (١٠,٢٪) | لغة الكشف |
| ٣٩٣ (١٢,٩٪) | ٢٩٨ (٣٧,٠٪) | الكشف |
| ٩٨٣ (٣٢,٤٪) | ٢٧٩ (٣٥,٠٪) | البحث |
| ٥٠٣ (١٦,٦٪) | ١٩٩ (٢٥,٠٪) | قصور التفاعل بين المستفيد والنظام |
| ٧٨ (٢,٥٪) | ١١ (١,٤٪) | أخرى |

الشكل رقم (٥٠) : الفئات الرئيسية للأخطاء في تقييم المدلرز (يزيد المجموع على ١٠٠ بالمئة نظراً لرد بعض الأخطاء إلى أكثر من سبب واحد).

دراسات فعالية التكلفة :

فعالية التكلفة هي العلاقة بين مستوى الكفاءة (الفعالية) والتكلفة التي ينطوي عليها تحقيق هذا المستوى. وربما كان من الممكن اتباع عدة طرق لبلوغ مستوى كفاءة معين، كما يمكن حساب تكلفة هذه الطرق. أما عائد التكلفة فيبدل على العلاقة بين

عائدات ناتج أو خدمة معينة وتكاليف تقديم هذه الخدمة أو هذا الناتج. وقياس العائدات أصعب بوجه عام من قياس الكفاءة، وإن كانت العائدات بالمفهوم التجاري تساوي مردود الاستثمار. أما التعبير عائدات تكلفة الأداء فيدل على مجمل العلاقة بين كل من التكاليف ومستوى الفعالية والعائدات.

ومن الممكن قياس تكلفة نظام استرجاع المعلومات بناء على مدخلات الموارد (التمويل). وينبغي أخذ كل من التكاليف الثابتة والتكاليف المتغيرة في الحسبان. وتشمل التكاليف الثابتة شراء التجهيزات أو استئجارها، وتكاليف التطوير، والتكاليف التي ينطوي عليها الاقتناء والتكشيف في مرصد البيانات الحالي. وهناك نوعان من التكاليف المتغيرة: النوع الأول يتوقف على عدد الواقعات؛ فإذا كان عدد عمليات البحث التي يتم إجراؤها قد ازداد، على سبيل المثال، من ١٠٠٠ عملية إلى ١٥٠٠ سنوياً، فإن تكلفة عملية البحث يمكن أن تنخفض س دولار. أما النوع الثاني من التكاليف المتغيرة فيتوقف على الطرق المختلفة لتشغيل النظام. فمن الممكن على سبيل المثال تغيير تكلفة البحث الراجع بتغيير طريقة التفاعل مع المستفيد (الزيارة الشخصية أو الهاتف) أو بإضافة أو إلغاء عملية غربلة المخرجات، أو بتغيير المستوى المهني لمن يقومون بإجراء عمليات البحث.

ومن بين العوامل التي تحول دون تطبيق تحليل فعالية التكلفة أو عائد التكلفة على نظم الاسترجاع، أن حساب التكلفة الحقيقية لمنتجات المعلومات وخدماتها ليس بالمهمة الهينة، ونادراً ما تنشر بيانات التكلفة المتصلة بمختلف الخدمات. ومن الممكن أن نقرر بسهولة ضغط النفقات بشكل ما، بخفض تكلفة التكشيف، على سبيل المثال، وذلك بإلغاء بعض المواد أو بالحد من الشمول في التكشيف، ولكن كيف يتسنى لنا قياس فعالية تكلفة الارتفاع بمتوسط كفاءة النظام من ٦٠ بالمئة استدعاء عند ٣٠ بالمئة تحقيق إلى ٦٠ بالمئة استدعاء عند ٥٠ بالمئة تحقيق مثلاً؟

وقد ناقش عدة مؤلفين مشكلات حساب تكاليف نظم استرجاع المعلومات. وقد انتهى مارون (Marron 1969) إلى أنه لا يمكن تطبيق الإجراءات المعيارية لحساب التكاليف، كما تستعمل في التطبيقات التقليدية لإدارة الأعمال، على خدمات مراكز المعلومات؛ فمرصد البيانات ليس منظاراً للمعدات أو الأجهزة وخاصة فيما يتصل

بتناقص القيمة أو الاستهلاك. ومن الصعب بوجه خاص توزيع التكاليف في حالة اعتماد عدد كبير من المخرجات المختلفة، كنشرة المستخلصات مثلاً، ونشرة الإحاطة الجارية المعتمدة على العناوين، والبحث الراجع، على نفس مرصد البيانات. ومن السهل حساب تكاليف المخرجات المباشرة، ولكن كيف يتسنى توزيع تكاليف المدخلات على مختلف المنتجات أو الخدمات؟ وقد تعرض مارون لهذه المشكلة إلا أنه لا يقدم حلاً مرضياً تماماً.

وإذا افترضنا أنه من الممكن توزيع تكاليف المدخلات بشكل واقعي، فإنه يمكن أن يكون من الممكن الخروج بسلسلة من تكاليف الوحدات لمختلف المنتجات والخدمات الوراقية. وعادة ما تشمل تكاليف الوحدة تكلفة كل واقعة بحث راجع، وتكلفة كل تسجيلية مسترجعة، وكل تسجيلية مطبوعة، وكل صفحة مطبوعة، أو كل مادة يتم بثها. وتعتبر هذه التكاليف عن تغيرات في الكم وليس تغيرات في مستوى الكفاءة. فانظر إلى تكلفة التسجيلية المسترجعة، والتي تبلغ ٧٤ سنتاً مثلاً، حيث تختلف تكلفة الوحدة هذه تبعاً لحجم المخرجات؛ فإذا تغيرت استراتيجيات البحث بحيث يسترجع كل بحث الآن، في المتوسط، ضعف ما كان يسترجعه من تسجيليات، فإن تكلفة كل تسجيلية مسترجعة يمكن أن تنخفض بشكل ملحوظ، إلى ٤٠ سنتاً مثلاً (فلا يمكن أن تنخفض إلى النصف نظراً لأن الزيادة في حجم ناتج البحث سوف ترتبط بها بعض التكاليف الإضافية، في وقت الاتصال على الخط المباشر مثلاً أو في تكلفة الطباعة). إلا أنه من الممكن للتغيير في استراتيجيات البحث أن يؤدي إلى انخفاض حاد في الكفاءة؛ حيث تؤدي توسعة البحث إلى زيادة في عدد الوثائق غير الصالحة، وبذلك يمكن لمتوسط التحقيق في البحث أن ينخفض من ٥٠ بالمئة إلى ٢٥ بالمئة.

ومما لا شك فيه، أنه لأغراض فعالية التكلفة، ينبغي أن تكون تكاليف الوحدة حساسة لما يطرأ على فعالية النظام من تغيرات. وتكلفة الوحدة بالنسبة للتسجيلية الصالحة المسترجعة (ج ر) من هذه النوعية من التكاليف. انظر إلى نظام يعمل على مستوى (ج ر) حوالي ١,٤٨ دولار، وعندما أدخلت بعض التغيرات على النظام لرفع متوسط الأداء في الاستدعاء أو التحقيق، فإن (ج ر) يمكن أن تنخفض. ومن ثم، فإن (ج ر) وحدة مفيدة للتعبير عما يمكن أن يطرأ على فعالية تكلفة النظام من تحسن.

ومن الممكن استعمال (ج ر) للمقارنة بين مختلف طرق التشغيل في النظام نفسه، أو المقارنة بين فعالية تكلفة نظامين مختلفين أو أكثر. ومن ثم فإنه يمكن لهذا المقياس تقييم الآثار الاقتصادية للتغيرات التي يمكن أن تدخل على إجراءات التكشيف، أو لغة التكشيف، أو استراتيجيات البحث، أو طريقة التفاعل مع المستفيد. فمن الممكن على سبيل المثال لرفع متوسط الشمول في التكشيف أن يسفر عن تحسين جوهري في الاستدعاء. وحتى على الرغم من زيادة تكاليف التكشيف، فإن ذلك يمكن أن يعني انخفاضاً ملحوظاً في (ج ر). وبذلك يمكن استعمال (ج ر) لتقييم فعالية تكلفة كل من البحث المفوض والبحث غير المفوض.

ولتحليل فعالية التكلفة أهميته بالنسبة لدراسة مدى توزيع الموارد المتاحة، بما يكفل تحقيق أقصى عائد ممكن للاستثمارات الموجهة، في الوثائق الصالحة المسترجعة مثلاً. فمن الممكن بلوغ الحد الأقصى لفعالية التكلفة عندما يحقق كل دولار من الموارد أفضل عائد ممكن، ممثلاً في زيادة الخدمات أو الارتفاع بمستواها. وعادة ما يتم إجراء تحليل فعالية التكلفة للتعرف على الطريقة الأقل تكلفة من بين عدة طرق مختلفة لبلوغ مستوى معين للخدمة. ومن الممكن تحسين فعالية تكلفة الخدمة، إما بالمحافظة على مستوى الكفاءة الحالي ولكن مع الحد من تكلفة تحقيقه، أو بجعل التكاليف ثابتة مع الارتفاع بمستوى الكفاءة. كذلك يمكن تحسين فعالية التكلفة بالارتفاع بمستوى الكفاءة، مع الحد في الوقت نفسه من التكاليف. إلا أنه ما لم تكن البداية من موقف سيء بشكل غير عادي، فإن هذا الضرب من التحسن نادراً ما يكون ممكناً.

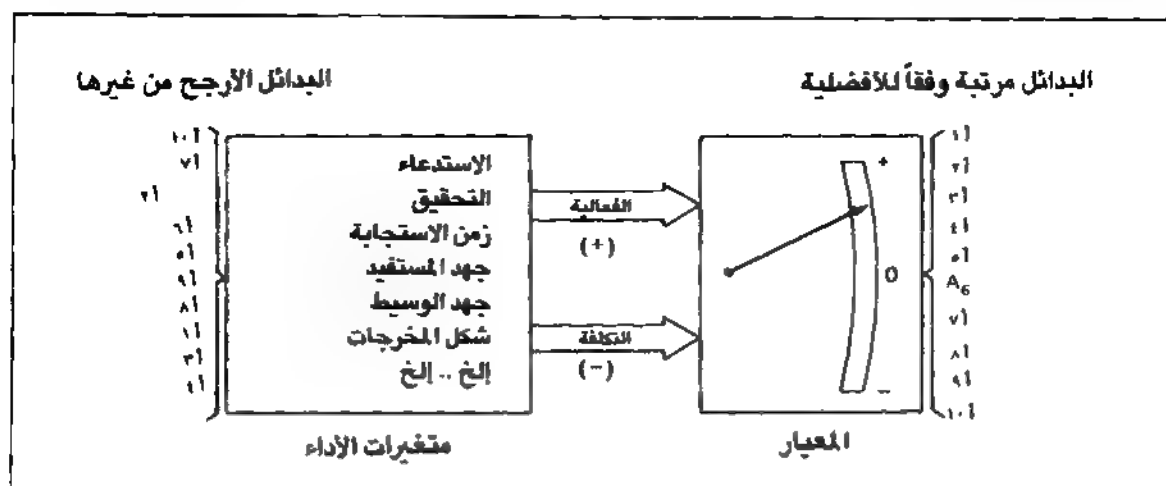
ووفقاً لما ذهب إليه هتش وماكين (Hitch and McKean) (1960) فإن تحليل فعالية التكلفة ينطوي على خمس خطوات أساسية :

- ١ - تحديد الأهداف التي يتعين بلوغها .
- ٢ - التعرف على الطرق المختلفة لتحقيق الأهداف .
- ٣ - التعرف على تكاليف مختلف الطرق البديلة .
- ٤ - وضع النماذج (كالمعادلات الرياضية، أو برنامج للحاسب الآلي، أو مجرد وصف لفظي مكتمل للموقف) التي تربط تكاليف كل بديل بالمدى الذي يؤدي به إلى تحقيق الأهداف.

٥ - وضع معيار لترتيب البدائل طبقياً وفقاً لأفضلياتها. ويكفل هذا المعيار طريقة لوزن التكاليف المقدرة مقابل الفعالية المقدرة

ويوضح الشكل رقم (٥١) بنية برنامج تحليل فعالية التكلفة . وقد تم التحقق من سلسلة من البدائل المحتملة للخدمة، تنطوي على مختلف التجميعات المؤتلفة للنتائج المتصلة بمختلف معايير الأداء (وبعضها ، ٢أ ، ٨أ ، ٢أ ، يمكن اعتباره أرجح كفة من غيره لأول وهلة) . ويتم تقييم أثر هذه التجمعات المؤتلفة على التكلفة وعلى فعالية التكلفة ، بما يكفل ترتيب بدائل الخدمة طبقاً.

كيف يمكن لتحليل فعالية التكلفة أن يطبق في المقارنة بين الاستراتيجيات البديلة لبلوغ نتائج مستهدفة معينة ؟ فيمكن لإحدى الشركات أن ترغب في معرفة ما إذا كان من الأفضل من وجهة نظر فعالية التكلفة إجراء جميع عمليات البحث في الإنتاج الفكري عن طريق المكتبة، أم بتدريب الباحثين العلميين بالشركة على إجراء عمليات البحث الخاصة بهم. والخطوة الأولى هي التحقق من جميع عناصر التكلفة التي تنطوي عليها المقارنة.



الشكل رقم (٥١) : بنية برنامج تحليل فعالية التكلفة. هناك تفاوت في تأثير مختلف التجمعات المؤلفة من متغيرات الاداء على التكلفة وفعالية التكلفة الخاصة بالنظام، بما يكفل تقييم بدائل الخدمة الأرحم من غيرها، وفقاً لمعيار فعالية التكلفة.

نفترض أن تكلفة الساعة في استجواب مراصد البيانات التي تفيد منها الشركة، كانت في المتوسط ٨٠ دولاراً (لا تشمل تكلفة القوى البشرية) ، وأن المكتبي يستنفد في المتوسط ١٥ دقيقة على الخط المباشر ، لكل عملية بحث واحدة ، أما الباحث العلمي فيستنفد ٢٠ دقيقة . وتكلفة المكتبي التي تتحملها الشركة ٢٠ دولاراً في الساعة ، بينما

تكلفة الباحث العلمي ٣٠ دولاراً في الساعة . وإذا وضعنا هذه العوامل في الاعتبار ، فإنه يتبين أن متوسط تكلفة عملية البحث التي يجريها الباحث العلمي حوالي ٣٧ دولاراً ($80 \text{ دولار} / \text{ساعة} + 30 \text{ دولار} / \text{ساعة} \times [1 / 2 \text{ ساعة}]$) بينما متوسط تكلفة عملية البحث التي يجريها المكتبي ٢٥ دولاراً ($80 \text{ دولار} / \text{ساعة} + 20 \text{ دولار} / \text{ساعة} \times [1 / 4 \text{ ساعة}]$) . إلا أن هذه المقارنة تغفل التكاليف المرتبطة بتفويض عملية البحث من جانب الباحث العلمي إلى المكتبي . فلكي يصف الباحث العلمي حاجته إلى المعلومات للمكتبي ، فإن ذلك قد يستنفد ١٥ دقيقة ، ومن ثم فإنه بالنسبة للموقف الخاص بالبحث المفوض ، يتعين إضافة ٧,٥٠ دولارات من وقت الباحث العلمي (٣٠ دولاراً / ساعة $\times [1 / 4 \text{ ساعة}]$) و ٥ دولارات من وقت المكتبي (٢٠ دولاراً / ساعة $\times [1 / 4 \text{ ساعة}]$) ، وبذلك ترتفع التكلفة إلى ٣٧,٥ دولاراً . وحتى هذا التحليل يفتقر للاكتمال ؛ فإذا افترضنا أن المكتبي قد تدرب فعلاً على البحث على الخط المباشر ، بينما الباحث العلمي لم يتدرب ، فإنه سوف يكون من الضروري وضع تكلفة تدريب الباحث العلمي في الحسابات ، كما يتعين استهلاك هذه التكلفة على فترة زمنية معينة .

وهذه الأرقام افتراضية بحتة ، ويستند التحليل إلى افتراضات معينة (أن جميع الباحثين العلميين ، مثلاً ، لديهم منافذ ملائمة في مكاتبهم) . وقد تم تبسيط هذا المثال عمداً لتوضيح ما يمكن أن يحدث عند إغفال عناصر تكلفة مهمة . ويمكن لتحليل أكثر تعقداً أن يضع في اعتباره كيفية قضاء المكتبيين والباحثين العلميين لوقتهم ما لم يكونوا يبحثون على الخط المباشر .

ومن الممكن المقارنة بين البحث على الخط المباشر بواسطة المكتبيين والبحث بواسطة الباحثين العلميين ، بناءً على التكلفة ، والفعالية ، وفعالية التكلفة . لننظر في البيانات التالية :

| الباحث العلمي | المكتبي | |
|-----------------------------------|------------------------------------|----------------|
| ٤٠ دولاراً لعملية البحث الواحدة | ٣٧ دولاراً لعملية البحث الواحدة | التكلفة |
| استرجاع ٢٠ وثيقة صالحة في المتوسط | استرجاع ١٥ وثيقة صالحة في المتوسط | الفعالية |
| ٢ دولار لكل وثيقة صالحة مسترجعة | ٢,٤٧ دولار لكل وثيقة صالحة مسترجعة | فعالية التكلفة |

فمقارنة التكلفة في صالح المكتبي ، أما مقارنة كل من الفعالية وفعالية التكلفة ففي صالح الباحث العلمي . ويمكن لتحليل أكثر تعقداً أن يميز بين الوثائق ذات الأهمية الكبرى والوثائق ذات الأهمية الدنيا ؛ حيث يمكن للباحث العلمي على سبيل المثال ، أن يعثر على وثائق أكثر من تلك التي يعثر عليها المكتبي ، إلا أنه من الممكن للمكتبي أن يعثر على الوثائق التي يراها الباحث العلمي الأكثر أهمية على الإطلاق ، وربما حدث ذلك لأن هذه الوثائق جديدة عليه .

ولم ينظر سوى عدد قليل جداً من الدراسات في فعالية التكلفة الحقيقية لإجراء البحث بواسطة المستفيد النهائي ، في مقابل البحث بواسطة المكتبي ، وإن كان كل من نيكولاس وإرباخ وهارس (Nicholas, Erbach and Harris (1987 قد تعرضوا لبعض القضايا . وهناك تحليل مماثل ينطوي على المقارنة بين فعالية تكلفة عمليات البحث في مرصد البيانات على الخط المباشر وفعالية تكلفة عمليات البحث عن الموضوعات نفسها في الكشافات المطبوعة (أو في مرصد البيانات في شكل الأسطوانات الضوئية المكتتزة) . وفي مثل هذا الموقف يتعين أيضاً الحرص على أخذ جميع تكاليف كل من الخيارين في الحسبان . فعند مقارنة تكلفة البحث في أداة مطبوعة كالـ *Chemical Abstracts* مثلاً ، بتكلفة البحث في نظيرها المتاح على الخط المباشر ، فإنه ينبغي أخذ تكلفة اقتناء الأداة المطبوعة في الحسبان . ومن بين العناصر الأساسية في تكلفة البحث على الخط المباشر تكلفة التعامل مع مرصد البيانات ، بما في ذلك تكلفة الحاسب الآلي ، وتكلفة الاتصالات بعيدة المدى ، وحقوق التأليف الخاصة بمرصد البيانات . أما التكلفة المناظرة الخاصة بالتعامل مع الأداة المطبوعة فهي تكلفة الاشتراك ، وتكلفة تداولها (من إجراءات الاشتراك والتسجيل والمراجعة والمطالبة بالأعداد المفقودة ... إلخ) وتكلفة الحيز الذي تشغله . وعلى ذلك ، فإنه إذا كانت التكلفة السنوية لاقتناء مرصد بيانات معين في شكله المطبوع ٥٠٠٠ دولار ، وكان يستعمل ٥٠٠ مرة في العام ، فإن تكلفة إجراء عملية البحث الواحدة ينبغي أن تشمل ١٠ دولارات مقابل « تكلفة التعامل » . وما لم يتم ذلك فإنه يمكن الخروج بصورة مشوهة تماماً في المقارنة بين البحث اليدوي والبحث على الخط المباشر (Elchesen, 1978; Lancaster, 1981) .

وقد وضع كل من هوانج وماكهيل (Huang and McHale (1990 نموذجاً لفعالية

التكلفة، لمساعدة المكتبيين في اتخاذ قرار متى يتوقف الاشتراك في مصدر مطبوع ليتم الاعتماد كلية على التعامل مع ذلك المصدر على الخط المباشر. ويربط « حد المطبوع / على الخط المباشر » تكلفة إتاحة المصدر المطبوع بالمكتبة، بمتوسط تكلفة إجراء البحث على الخط المباشر في مرصد البيانات المناظر. ويتم الاعتماد على متوسط التكلفة السنوية للمصدر المطبوع (معدل الاشتراك السنوي) لحساب متوسط التكلفة اليومية ، وهي عبارة عن تكلفة الاشتراك مقسومة على عدد أيام فتح المكتبة في العام . فإذا كان متوسط تكلفة البحث على الخط المباشر مساوياً لهذه التكلفة اليومية أو أقل منها ، فإنه من المفترض توقف الاشتراك في المصدر المطبوع . وهذا الأسلوب في التحليل غير مسبوق إلا أنه تبسيطي ؛ فمن الصعب إدراك السبب في اتخاذ متوسط التكلفة اليومية بديلاً عن تكلفة واقعة الإفادة من المصدر المطبوع ، اللهم إلا الحاجة إلى إجراء دراسة وصفية تحليلية لتقدير الإفادة السنوية ، بينما يمكن حساب متوسط التكلفة اليومية بسهولة (مع مراعاة أن التكلفة الحقيقية للاقتناء تتضمن الحيز الذي يشغله المصدر المطبوع، وتداول هذا المصدر وعناصر تكلفة أخرى ، فضلاً عن تكلفة الاشتراك) . فالنموذج في الواقع ليس نموذجاً حقيقياً لفعالية التكلفة ، لأنه لم يأخذ في حسابه فعالية البحث (أي أنه يفترض تساوي كل من عمليات البحث في المصدر المطبوع والبحث على الخط المباشر، في الفعالية) .

ويهتم المكتبيون الآن بالبحث في المصادر المطبوعة مقابل البحث في الأسطوانات الضوئية المكتنزة ، والبحث في الأسطوانات الضوئية المكتنزة في مقابل الخط المباشر ، وكذلك البحث في المصادر المطبوعة مقابل البحث على الخط المباشر. ويقدم ولش Welsh (1989) مثالاً للمقارنة بين الخط المباشر والأسطوانات الضوئية المكتنزة ، تعتمد على الإفادة من مرصد بيانات NTIS ، والتي تقدر بـ ١٦٢ عملية بحث أو ٦٤ ساعة سنوياً في مكتبته. ويقدر ولش تكلفة الساعة بالنسبة للأسطوانات الضوئية المكتنزة بـ ٣٥,١٦ دولاراً (تكلفة الاشتراك السنوي في مرصد البيانات ٢٢٥٠ دولاراً مقسمة على ٦٤) ، وذلك في مقابل ٨٠ دولاراً تكلفة الساعة بالنسبة للتعامل مع DIALOG/DIALNET . وبمعدل ٦٤ ساعة بحث سنوياً يبلغ الاقتصاد السنوي نتيجة اقتناء الأسطوانات الضوئية المكتنزة حوالي ٢٨٧٠ دولاراً (أي ٥١٢٠ دولاراً ناقص ٢٢٥٠ دولاراً) .

وكما يعترف ولش ، فإن هذه مقارنة تبسيطية للتكلفة ، فلم يؤخذ في الحساب بالنسبة للاتصال على الخط المباشر تكاليف طباعة التسجيلات الوراقية (٠,٣٠ دولار على الخط المباشر ، و ٠,٤٥ دولار خارج الخط المباشر) ، والتي يمكن أن تكون عنصراً مهماً في التكلفة الإجمالية للبحث الراجع الشامل . إلا أنه فيما يتصل بالأسطوانات الضوئية المكتنزة ، ينبغي مراعاة تكلفة الورق المستهلك . وما هو أهم من ذلك ، فإنه ينبغي تخصيص جزء ما من تكلفة اقتناء أجهزة الأسطوانات الضوئية المكتنزة لكل ساعة يتم فيها الإفادة من هذه الأسطوانات . فإذا افترضنا أن الأجهزة (محطة عمل وجهاز لتشغيل الأسطوانات الضوئية المكتنزة) تبلغ تكلفتها ٢١٩٥ دولاراً ، وأن العمر الافتراضي لهذه الأجهزة خمس سنوات ، وأنها تستعمل لمدة ١٦٠٠ ساعة للبحث خلال فترة السنوات الخمس (يستند هذا التقدير إلى خمسة مراصد بيانات على أسطوانات ضوئية مكتنزة ، يستعمل كل منها في المتوسط لمدة ٦٤ ساعة سنوياً) . ولهذا فإنه ينبغي إضافة حوالي ١,٣٧ دولار (٢١٩٥ دولاراً ÷ ١٦٠٠ ساعة) لتكلفة ساعة البحث في الأسطوانات الضوئية المكتنزة مقابل استخدام الأجهزة ، بالإضافة إلى ما يزيد على ذلك قليلاً مقابل الورق المستهلك والحيز الذي تشغله الأجهزة بالمكتبة (الذي يمكن أن يكون مناظراً تقريباً لمحطة العمل الخاصة بالأسطوانات ومنفذ الاتصال على الخط المباشر) . وعلى ذلك ، فإن التكلفة الفعلية لساعة البحث في الأسطوانات الضوئية المكتنزة يمكن أن تكون أقرب إلى ٣٧ دولاراً منها إلى ٣٥ دولاراً كما قدرها ولش ، على الرغم من أن هذا التقدير لا يزال أقل بكثير من تكلفة البحث على الخط المباشر .

إلا أن هذا التحليل يستند فقط إلى تكاليف التعامل مع مرصد البيانات ويتجاهل العنصر بالغ الأهمية الخاص بتكاليف العنصر البشري . فمن وجهة نظر المكتبة ، يتمتع مرصد البيانات المسجل على أسطوانات ضوئية مكتنزة بميزة إضافية ، وهي قدرة معظم المستفيدين من المكتبة على إجراء عمليات البحث الخاصة بهم ، بينما كانت عمليات البحث على الخط المباشر ، في مكتبة ولش (بأحد الأجهزة الحكومية) تتم بواسطة المكتبيين المهنيين . إلا أنه من وجهة نظر الجهاز يمكن للموقف أن يكون مختلفاً تمام الاختلاف ، حيث يمكن أن تكون مرتبات المستفيدين الذين يبحثون في مراصد البيانات المسجلة على أسطوانات ضوئية مكتنزة ، أعلى في المتوسط من مرتبات المكتبيين ، كما أنه

من المحتمل أن يستنفدوا في البحث وقتاً أطول مما يمكن أن يستنفده المكتبيون (وقد بين ولس [1989] Welsh فعلاً أن المستفيدين من الأسطوانات الضوئية المكتنزة يميلون لاستنفاد مزيد من الوقت في البحث لأنهم يعرفون أنهم لا يدفعون مقابل وقت الاتصال)، ومن ثم، فإن التكلفة الفعلية لكل بحث بالنسبة للجهاز، أخذين المرتبات والتكاليف غير المباشرة في الحسبان، يمكن أن تكون أعلى بكثير بالنسبة للتعامل مع الأسطوانات الضوئية المكتنزة، مما هي عليه بالنسبة للخط المباشر.

ولا تأخذ هذه المقارنة في الحسبان سوى جانب التكلفة لمعادلة فعالية التكلفة، أو على الأقل، تأخذ في الحسبان تكلفة عملية البحث باعتبارها وحدة فعالية التكلفة، لا تكلفة كل وثيقة صالحة مسترجعة. فإذا كان بإمكان المكتبي العثور على وثائق صالحة عن طريق إمكانات الخط المباشر، أكثر بكثير مما يمكن للمستفيدين من المكتبة العثور عليه في مراصد البيانات المسجلة على أسطوانات ضوئية مكتنزة، فإن تكلفة الوثيقة الصالحة المسترجعة (المقياس الحقيقي لفعالية التكلفة) يمكن أن تكون بالنسبة للبحث على الخط المباشر أقل مما هي عليه بالنسبة للأسطوانات الضوئية المكتنزة. ومن ناحية أخرى يمكن من وجهة نظر الجهاز أن يكون البديل الأفضل بناء على فعالية التكلفة هو البديل الذي يقوم فيه المكتبيون بإجراء عمليات البحث في الأسطوانات الضوئية المكتنزة لصالح المستفيدين من المكتبة. ومما لا شك فيه أن هذه المقارنة معقدة جداً، حيث لا يمكن لقرار تحديد البديل الأفضل أن يستند إلى التكاليف دون سواها، وإنما ينبغي أن يأخذ نتائج عملية البحث (الفعالية) في الحسبان. هذا بالإضافة إلى أن القرار يتوقف وإلى حد بعيد على ما إذا كان إجمالي تكاليف الجهاز يؤخذ في الحسبان أم تكاليف المكتبة فقط.

كذلك يمكن تطبيق تحليل فعالية التكلفة على علاقات المدخلات بالمخرجات. وتنظر ماندل (1988) Mandel إلى أحد المظاهر التي تتبدى فيها العلاقة بين تكاليف المدخلات وكفاءة البحث؛ فهي تربط مختلف مستويات التفصيل في الفهرسة بـ :

- (١) احتمال بحث المستفيدين بناء على ما يشتمل عليه الفهرس من نقاط الوصول.
- (٢) العدد المحتمل لعمليات البحث التي يمكن أن تكون ناجحة تبعاً لاختلاف مستويات التفصيل.

وتهتم دراسات فعالية التكلفة أكثر ما تهتم بالبدائل والعائدات المتناقصة ، أي تحديد المستوى الذي يمكن عنده لنظام الاسترجاع أن يعمل بكفاءة وبشكل اقتصادي. وقد سبق أن تعرضنا لأمثلة لمثل هذه الاعتبارات في أماكن أخرى من هذا الكتاب. وهي تشمل القرارات المتعلقة بما يمكن تغطيته في مرصد البيانات (فمن الممكن على سبيل المثال أن يكون هناك حد أقصى لمجموعة الدوريات التي يمكن تكسيها لتلبية احتياجات المستفيدين من مرصد البيانات) . وما مدى الشمول في تكشيف الوثائق على اختلاف أنواعها في مرصد البيانات. وترتبط ظاهرة تناقص العائدات ارتباطاً وثيقاً بظاهرة القوانين القياسوراقية وما يتصل بها (قانون برادفورد (1953) Bradford وقانون زيف Zipf (1935) بوجه خاص) ، وبفكرة « مكتبة التسعين بالمئة » (Bourne, 1965) التي تقر بأنه يمكن لمرفق المعلومات تلبية نسبة مئوية معينة من احتياجات المستفيدين (ربما تصل إلى ٩٠ بالمئة) بكفاءة وبشكل اقتصادي ، إلا أنها لكي ترتفع عن مستوى الأداء هذا يمكن أن تحتاج إلى مستوى للاستثمار غير متناسب على الإطلاق. فالاحتياجات التي يمكن تلبيتها بكفاءة تميل لأن تكون متجانسة ويمكن التنبؤ بها. وهناك جوانب أخرى لتحليل فعالية التكلفة وعائد التكلفة تناولها كل من لانكستر (1993 a) Lancaster وفلاوردو وهوايتهيد (1974) Flowerdew and Whitehead .

ويرتبط أوسع أنواع دراسات فعالية التكلفة بالحد الأقصى لما يمكن توزيعه من الموارد المالية للمؤسسة على كل ما تقدمه من خدمات. وتهدف هذه الدراسات إلى معرفة ما إذا كان من الممكن لتوزيع آخر للموارد أن يسفر عن تحسن في الخدمات. والمشكلة الأساسية هي تحديد أفضليات المستفيدين وأولوياتهم بالنسبة لمختلف الخدمات، بحيث يمكن توزيع الميزانية بما يتفق وأولويات المستفيدين . ويصف رافل وششكو (1969) Raffel and Shishko أسلوباً مفيداً لتحقيق هذا التوزيع . وينطوي هذا الأسلوب، بإيجاز ، على « مباراة إدارية » يشارك فيها عينات من مجتمع المستفيدين ، حيث يقدم للمستفيدين قائمة بالخدمات المحتملة والميزانيات المحتملة المتعددة ، ومن بينها الميزانية القائمة للمؤسسة . كما يقدم إليهم أيضاً قائمة بالنتائج المحتملة (العائدات والجزاءات) المترتبة على حدوث تغييرات في أنماط الخدمة ، والتكاليف المرتبطة بمختلف البدائل. ويطلب منهم توزيع الموارد المرصودة على مختلف احتمالات الخدمات للتعبير عن أفضلياتهم واحتياجاتهم . ويعطي توزيع الموارد على هذا النحو

لمركز المعلومات فكرة عن احتياجات المستفيدين أوضح مما يمكن الخروج به من مجرد سؤال المستفيدين عن أفضلياتهم على نحو أكثر تجريداً ، كما يكفل أيضاً التحقق من مجموعات مختلفة من الأولويات الخاصة بمختلف قطاعات مجتمع المستفيدين. ومن المفيد أيضاً للإدارة أن تدرك كيف يمكن لأفضليات المستفيدين أن تتغير تبعاً لزيادة الميزانية أو نقصانها.

تحليل عائد التكلفة :

تحاول دراسة عائد التكلفة الربط بين تكلفة تقديم خدمة ما بعائدات توافر هذه الخدمة. فالخدمة تجد ما يبررها إذا زادت العائدات عن التكاليف . وللأسف ، فإنه على الرغم من أن مبدأ تحليل عائد التكلفة يبدو بسيطاً ، فإنه ليس من السهل إجراء مثل هذه الدراسات ، نظراً لصعوبات قياس عائدات نظم استرجاع المعلومات، وخاصة إسباغ قيم مالية على هذه العائدات. وتشمل بعض المعايير المحتملة لتحديد نسب عائد التكلفة لنظم استرجاع المعلومات:

- ١ - الاقتصاد المحتمل في التكاليف عن طريق الإفادة من الخدمات، بالمقارنة بتكاليف الحصول على المعلومات أو الوثائق اللازمة من مصادر أخرى.
 - ٢ - تجنب انخفاض الإنتاجية ، إنتاجية الطلبة مثلاً ، أو أعضاء هيئة التدريس أو الباحثين ، ذلك الانخفاض الذي كان يمكن أن يترتب على عدم توافر مصادر المعلومات في متناول هؤلاء .
 - ٣ - تحسين عملية اتخاذ القرارات ، أو خفض مستوى القوى البشرية اللازمة لاتخاذ القرارات .
 - ٤ - تجنب التكرار أو الفاقد في جهود البحث والتطوير في المشروعات التي سبق تنفيذها أو تلك التي أثبتت الدراسات السابقة عدم جدواها .
 - ٥ - الحث على الابتكار أو الإنتاجية بإتاحة الإنتاج الفكري المتصل بالتطورات الجارية في مجال معين.
- وقد حاول بعض المكتبيين العاملين في قطاع الصناعة تبرير وجود خدماتهم

بحساب كم ما كان يمكن أن تتكبده الشركات من تكاليف لشراء الخدمات المناظرة من أي مكان آخر ، وليكن مكتبة أخرى في المؤسسة ، أو مكتبة خارجية ، أو مرفق معلومات تجارياً (انظر على سبيل المثال Magson, 1973). والافتراض الأساس في هذه الحالة هو أن الخدمة جديرة فعلاً بالاهتمام ، أما السؤال الذي يؤخذ في الحسبان ، فهو ما إذا كان من الأفضل للشركة تقديم الخدمة عن طريق مكتبة محلية أم بطريقة أخرى^(١). ويمكن لجميع ما تقدمه المكتبة من خدمات ، أو خدمة واحدة فقط بعينها ، أن تكون بؤرة الاهتمام؛ فمن الممكن أن يتبين على سبيل المثال ، أن إجمالي تكلفة تقديم خدمات البحث في الإنتاج الفكري على الخط المباشر ، عن طريق المكتبة ، وعند مستوى ٥٠٠ عملية بحث في العام ٣٥٠٠٠ دولار. ويمكن للحصول على هذا المستوى من الخدمة من مؤسسة تجارية أن تبلغ تكلفته ٥٠٠٠٠ دولار. إذن فالعائد الصافي للشركة من تقديم الخدمة محلياً يبلغ ١٥٠٠٠ دولار.

وهذا هو السبيل المناسب ، على أكثر من نحو ، لتبرير وجود خدمة معلومات من هذا النوع ، على الرغم من أن هناك بعض المشكلات التي ترتبط به ؛ فبعض الخدمات قد لا تكون مناسبة لأن يعهد بها إلى مقاول أو متعهد خارجي ، وذلك لدواعي الأمن العملي أو الصناعي. وقد تكون هناك إحدى المزايا المرتبطة بالإمكانات المحلية ، لا يستطيع المتعهد الخارجي ضمانها ، على الرغم من أنه قد يكون من الصعب إضفاء أي قيمة مالية حقيقية على هذا العامل. وهناك مشكلة ترتبط بالظروف العملية أكثر من غيرها. وهي أنه من الممكن لمجموعة مقتنيات المكتبة أن تسهم في كثير من الخدمات المختلفة ، كالإمداد بالوثائق ، والبحث في الإنتاج الفكري ، والرد على الاستفسارات ، وإعداد النشرة الإعلامية ، ومن الصعب توزيع تكاليف المقتنيات على مختلف الخدمات على نحو معقول. ورغم ذلك ، فإن إدارة المؤسسة إذا ما اقتنعت بهذا السبيل لتبرير الخدمات ، فإنها تتحمل عبئاً كبيراً للتوصية به من وجهة نظر المكتبيين.

وهناك أسلوب آخر مماثل يقارن تكلفة المكتبي الذي يقدم خدمة معينة بتكلفة المستفيدين من المكتبي الذين يقومون بالنشاط لأنفسهم (Mason, 1972; Rosenberg, 1969)

(١) ربما كان هناك من يرى أن هذه أقرب إلى دراسة فعالية التكلفة ، التي تقارن بين استراتيجيات بديلة ، منها إلى دراسة عائد التكلفة.

ونفترض، كمثال بسيط ، أن متوسط تكلفة بحث الإنتاج الفكري الذي يقوم به المكتبي ١٤٠ دولاراً ، وأن متوسط تكلفة البحث المناظر الذي يقوم به المستفيد ١٩٥ دولاراً ، نظراً لتفاوت المرتبات ، فإنه قد يكون من الممكن القول حينئذ بأن المكتبي يوفر للشركة ٥٥ دولاراً لكل عملية بحث يقوم بإجرائها. وهنا بعض الافتراضات الأساسية ؛ كأن يكون من الممكن للمستفيد أن يجري عملية البحث في حالة عدم تواجد المكتبي ، وأن يكون من الممكن للنتائج التي يتوصل إليها المستفيد أن تكون مناظرة من الناحية النوعية لتلك التي يتوصل إليها المكتبي.

وقد حاول روزنبرج (1969) Rosenberg التوسع في هذه الطريقة بأن يطلب من المستفيدين تقييم نتائج البحث في الإنتاج الفكري الذي يجريه المكتبي ، وفقاً للميزان التالي:

٠ لا قيمة له (غير صالح مثلاً أو وصل بعد فوات الأوان) .

١. ملائم ، كان من الممكن للمستفيد أن يستنفد الوقت نفسه المستنفد من قبل المكتبي.

٢. جيد ، كان من الممكن للمستفيد أن يستنفد ضعف الوقت الذي استنفده المكتبي.

٣. ممتاز ، ما كان يمكن للمستفيد أن يحقق النتائج ، أو ما كان له أن يحققها بتكلفة معقولة.

ومن الممكن وضع هذه الأوزان في معادلة « توفير » على النحو التالي:

$$(أ \times ب \times ج) - ١$$

حيث « أ » هي تكلفة وقت المكتبي ، و « ب » عامل ضرب لمراعاة الاختلاف في المرتب بين المكتبي والمستفيد ، و « ج » هو عامل الوزن الذي حصلت عليه نتائج عملية البحث. وعلى ذلك ، فإنه إذا كان البحث في الإنتاج الفكري يتكلف ٧٥ دولاراً من وقت المكتبي ، وكان التفاوت في المرتب ١,٥ ، والوزن الذي أعطاه المستفيد للنتائج ٢ ، فإنه من الممكن إذن حساب التوفير على النحو التالي :

$$(٧٥ \text{ دولاراً} \times ١,٥ \times ٢) - ٧٥ ، أي ١٥٠ دولاراً .$$

ومما لا شك فيه أن هذا أسلوب ذاتي جداً ، حيث لا يمكن أن نطمئن كثيراً إلى أن

بإمكان المستفيد التوصل إلى أي تقدير حقيقي لدى ما كان يمكن أن يستنفد من وقت لتنفيذ إحدى مهام استرجاع المعلومات. وقد حاول عدد آخر من الباحثين تبرير خدمة المعلومات بتقدير الاقتصاد المحتمل في التكلفة والذي يمكن أن يعزى إلى توافر هذه الخدمات. فمن المفترض أنه ما لم تتوافر المكتبة فإنه من الممكن للمهندسين أو الباحثين أن ينفقوا المزيد من وقتهم في البحث عن المعلومات ، وأن ذلك يمكن أن يكون مكلفاً بالنسبة للمؤسسة. وتتوقف صلاحية هذا النوع من التحليلات على مدى سلامة تقديرات الوقت الذي يمكن اقتصاده.

وقد حاول نايتنجيل (1973) Nightingale إجراء تحليل لعائد التكلفة لنشرة مستخلصات إحدى الشركات. فقد تبين له بالحساب أن تكلفة إنتاج هذه النشرة تبلغ ٢٥٠٠ دولار سنوياً ، واعتمد على المنهج الوصفي التحليلي لمعرفة عدد الدوريات التي كان يطلع عليها بانتظام من يتلقون النشرة ، كما طلب من القراء تقدير عدد الدوريات الإضافية التي يمكن لهم الاطلاع عليها إذا ما توقفت النشرة ، وانتهى إلى متوسط قيمته ست دوريات إضافية سنوياً . وتبين من حسابات نايتنجيل أن الاطلاع على الدورية يمكن أن يستنفد من المستفيد عشر دقائق في المتوسط للتحقق من المواد التي تهمة . وقد تم تقدير تكلفة هذا النشاط في حدود ١٨,٥ ساعة لكل مستفيد سنوياً (٦ دوريات × ١٠ دقائق × عدد إصدارات الدورية) ، بحوالي ٧٤ جنيهاً أسترلينياً . وعندما يكون هناك ٤٠٠ مستفيد ، فإن تكلفة الاطلاع الإضافي يمكن أن تكون ٧٤ جنيهاً × ٤٠٠ ، أي ٢٩٦٠٠ جنيه. فالحد الأدنى لتحليل عائد التكلفة إذن وفر سنوي قدره ٢٧١٠٠ جنيه (٢٩٦٠٠ ناقص الـ ٢٥٠٠ اللازمة لإنتاج النشرة) .

وتحليل نايتنجيل متحفظ بشكل معقول ، أما الأقل تحفظاً وبشكل واضح فتلك البيانات التي أوردها كرامر (1971) Kramer والمعتمدة على دراسة أجريت في شركة بوينج Boeing Aerospace : فقد قام كرامر بتقدير ما يتحقق للشركة من وفر نتيجة الاعتماد على المكتبة في إجراء عمليات البحث في الإنتاج الفكري والرد على أسئلة الحقائق، بدلاً من أن يقوم المهندسون بهذا العمل بأنفسهم. وبناء على الاستبيانات التي تمثل استجابة ١٥٣ مهندساً أجريت لهم عمليات بحث في الإنتاج الفكري ؛ تبين أن هناك ٩٤٧٩ ساعة من وقت المهندسين كان من الممكن أن تستنفد إذا قدر لهم إجراء عمليات

البحث بأنفسهم . وكان وقت المكتبيين المستنفد في إجراء عمليات البحث هذه يقدر بـ ١٠٧١ ساعة عمل (أي حوالي سبع ساعات عمل لكل عملية بحث) . ومما لا شك فيه، أنه حتى إذا كان المكتبي يتقاضى المقابل نفسه الذي يتقاضاه المهندس في الساعة، فإن الوفرة يمكن أن يكون كبيراً. فقد كان من الممكن لحوالي ٨٠٠٠ ساعة من وقت المهندس، بمعدلات عام ١٩٩٢ م (بما في ذلك جميع التكاليف غير المباشرة) أن تبلغ تكلفتها أكثر من ٣٠٠٠٠٠ دولار.

كذلك اعتمد كرامر (Kramer (1971 على إجراء مقابلات للمتابعة عبر الهاتف مع ٢١٥ مهندساً أجابت لهم المكتبة عن أسئلة حقائق . وبينما كان متوسط الوقت المستنفد في الإجابة عن السؤال من قبل المكتبيين ١٢ دقيقة، قرر المهندسون أنهم كانوا من الممكن أن يستنفدوا ٥,٤٢ ساعة في المتوسط للعثور على إجابة السؤال الواحد ! وبالنسبة لـ ٢١٥ سؤالاً فإن هذا يمثل وفراً آخر قدره ١١٦٦ ساعة من وقت المهندسين.

كذلك طلبت إستابروك (Estabrook (1986 من المهندسين تقدير الاقتصاد في الوقت المرتبط بالإفادة من إحدى خدمات الاسترجاع والإمداد بالوثائق، وأن يقدروا قيمة المعلومات المسترجعة بالدولار ، وانتهت من ذلك إلى أنه وفق أكثر التقديرات تحفظاً، اكتسبت الشركة دولارين مقابل كل دولار أنفقته على الخدمة. إلا أنه وفقاً لظروف الحالة المثل كان من الممكن للشركة أن تكتسب حوالي ٥٠ دولاراً مقابل كل دولار استثمرته في مركز المعلومات. وقد أمكن التوصل إلى هذه النتيجة الأخيرة بإدخال حالتين حديثتين قام فيهما من تلقوا المعلومات بتقدير العائد المحتمل بمليون دولار، ومليونين ونصف المليون من الدولارات.

وقد حاول أيضاً عدد آخر من الدراسات الإيعاز للمستفيدين بتحديد القيمة الدلالية للنتائج التي يتلقونها من إحدى خدمات المعلومات ؛ ويقدم كوليت وبريس (Collette and Price (1977 على سبيل المثال نموذجاً لتبرير تكلفة أنشطة البحث في الإنتاج الفكري على أساس ما يمكن إدخاره من وقت المهندسين وبعض العناصر الأخرى للقيمة المالية لنتائج عمليات البحث. واعتماداً على النتائج المستخلصة من دراسة وصفية تحليلية للمستفيدين ، انتهى كوليت وبريس إلى ما اعتبره تقديرات

«بالغة التحفظ» للعائد. فوفقاً للحساب بلغ عائد عملية البحث على الخط المباشر ٢١٥ دولاراً في المتوسط، بينما كانت تكلفة عملية البحث ١١٢ دولاراً في المتوسط. كذلك يؤكد كوليت وبراييس حقيقة غالباً ما يتم إغفالها، وهي أنه من الممكن حتى لعمليات البحث التي تجريها المكتبة ولا تسفر عن وثائق صالحة على الإطلاق، أن تكون لها قيمتها بالنسبة للشركة، نظراً لأنه قد يكون إجراء المهندسين لعمليات البحث هذه بأنفسهم أعلى تكلفة من إجرائها بواسطة المكتبة.

ومن أفضل دراسات عائدات عمليات البحث على الخط المباشر تلك الدراسة التي تناولها كل من جنسن وأسبري وكنج (Jensen, Asbury and King (1980)، حيث أجريت مقابلات هاتفية مع عينة من المستفيدين بمركز ناسا للتطبيقات الصناعية NASA Industrial Application Center بجامعة كاليفورنيا الجنوبية. فقد طلب من المستفيدين تقدير عائدات بحث أجري لهم على الخط المباشر، وذلك بناء على الساعات التي أمكن ادخارها (مقارناً بإجرائهم لعمليات البحث بأنفسهم، أو الحصول على ما يحتاجونه من معلومات من أي مصدر آخر) والقيمة المحتملة للمعلومات عند تطبيقها على ناتج قائم أو على إحدى العمليات أو الخدمات، أو على ناتج جديد مستهدف أو عملية أو خدمة جديدة. ومن بين المستفيدين الـ ١٥٩ الذين شملتهم الدراسة، كان بإمكان ٥٢ بالمئة التحقق من العائد بالدولار. وقد سجلت الدراسة أن العائدات الجارية تبلغ ٣٦٤٦٠٥ دولارات، بينما تبلغ العائدات بعد خمس سنوات ٨٧٣٥٠٠ دولار.

ويمكن لهذه النوعيات من البيانات أن تكفل تأييداً قوياً جداً لخدمات المعلومات، إذا كانت البيانات تتمتع بضمانات الثقة الكاملة. ومن الصعب للأسف الاطمئنان إلى أنه من الممكن لأي إنسان الخروج بتقدير واقعي لطول الوقت الذي يمكن أن يستنفد في إنجاز نشاط بعينه في استرجاع المعلومات، كما أن النتائج التي توصل إليها كرامر (Kramer (1971 من أن الوقت المستنفد من جانب المهندس للإجابة عن سؤال ما، أطول من الوقت المستنفد من جانب المكتبي بحوالي ٢٠ مرة، من الأمور التي تقوض مقومات الثقة. وربما كان من المستبعد أيضاً أن يكون بإمكان المستفيد من إحدى خدمات المعلومات تقدير القيمة المالية لما يتلقاه من معلومات، على أية درجة من الدقة، على الرغم من إدعاء إستابروك (Estabrook (1986 أن مثل هذه التقديرات يمكن أن تكون أفضل مما هو مفترض بوجه عام.

وحتى إذا لم يكن من الممكن تحديد القيمة المالية لمعلومات بعينها بدقة، فإنه من الممكن أن تكون هناك ظروف يمكن لخدمة المعلومات الصناعية أن تؤكد فيها جدارتها بالنسبة للمؤسسة على نحو درامي. وربما كان أعظم عائد يمكن للمكتبة أن تحققه بالنسبة لهيئات البحث هو الكشف عن معلومات تحول دون إجراء هذه الهيئات لبحوث أجريت في أي مكان آخر. ومن الصعب توثيق وقائع من هذا النوع (بل الأصعب من ذلك هو إثبات أنه ما كان للهيئة أن تعثر على المعلومات بدون المكتبة) إلا أنه من الممكن لحالة واحدة ، إذا ما تم توثيقها ، أن تبرر تكلفة المكتبة لعدة سنوات. وفي المملكة المتحدة أجرى مارتن (1964) Martyn دراسة ضخمة للتكرار غير المقصود للبحوث وتكلفة هذا التكرار، وانتهت هذه الدراسة إلى شواهد قوية تدعم القول بأنه من الممكن توفير مبالغ مالية ضخمة بالحرص على إجراء المزيد من عمليات البحث الشامل للإنتاج الفكري قبل البدء في تنفيذ مشروعات البحوث . ويقدم كوبر (1968) Cooper أرقاماً خاصة بالاقتصاد في وقت البحث الذي يمكن أن يعزى إلى الاتصالات غير الرسمية المتمثلة في الطباعات المبدئية ، التي تم توزيعها تجريبياً بواسطة جماعات تبادل المعلومات التي كونتها المعاهد القومية للصحة National Institutes of Health .

ومن بين المقاييس الأخرى المحتملة للعائد ، تناقص الإنتاجية الذي يمكن أن يترتب على عدم توافر المكتبة بالشركة ، مما يضطر الباحثين أو المهندسين للانتظار أوقاتاً طويلة لكي يحصلوا على ما يحتاجون من معلومات . فقد تبين للمر (1959) Mueller على سبيل المثال أن نشاط بعض المهندسين قد توقف فعلاً بينما هم ينتظرون المعلومات اللازمة لإنجاز إحدى المهام الحيوية. والافتراض هنا ، بالطبع ، هو أن الحصول على المعلومات يوفر الوقت . إلا أن سولومين (1974) Solomin يرى أنه في ظل ظروف معينة، يمكن لتوافر المعلومات أن يزيد مما تتكبده الشركة من تكاليف، نظراً للوقت المستنفد في تجهيز هذه المعلومات واستيعابها.

وقد نظر موندشاين (1990) Mondschein في العائدات الناتجة عن الإفادة من إمكانات البث الانتقائي للمعلومات في إحدى مؤسسات البحث العلمي، ولم تحظ التكاليف بالمناقشة . وكان المقياس المستخدم للعائد هو تحسين الإنتاجية مقيسة بالمطبوعات الناتجة. وقد اكتشف موندشاين أن المستفيدين المنتظمين من خدمات البث

الانتقائي للمعلومات أعلى إنتاجية من غير المستفيدين أو الباحثين الذين لا يفيدون من هذه الخدمات إلا لماماً.

وأخيراً يمكن للمكتبي أن يبين بعض ما يمكن أن يعود على الشركة من آثار إيجابية أخرى ، يمكن ردها إلى ما يقدمه المكتبيون من معلومات. وتشمل هذه الآثار تطوير المنتجات الجديدة ، والتعرف على سبل الحد من تكاليف المنتجات الحالية (ربما نتيجة لاستخدام مواد أقل سعراً إلا أنها لا تقل فعالية) أو الحصول على عقد مهم . وليس من السهل إثبات أن المكتبة كانت مسئولة مباشرة عن هذا النوع من الأحداث ، إلا أنه من الممكن لمثال واحد موثق أن يكون كافياً لتبرير وجود المكتبة في المستقبل .

وقد أجرى الباحثون بمؤسسة كنج للبحوث (King et al. (1982) تحليلاً لعائد التكلفة بمزيد من التعمق في محاولة للتعرف على أهمية مرصد بيانات الطاقة Energy Data Base الخاص بوزارة الطاقة بالولايات المتحدة الأمريكية (DOE) . فعن طريق استخدام الاستبيانات انتهى فريق البحث إلى أن الاطلاع على المقالات والتقارير من جانب العلماء والمهندسين الذين تمول وزارة الطاقة جهودهم قد أسفر عن الحصول على معلومات تحقق عائداً سنوياً قدره ١٢ بليون دولار ، (وذلك بتجنب تكرار الجهود والاقتصاد في الوقت وبطرق أخرى) . ويأتى ذلك في مقابل إنفاق وزارة الطاقة لـ ٥,٣ بلايين دولار سنوياً على البحث والتطوير ، و ٥٠٠ مليون دولار على تجهيز المعلومات والإفادة منها.

وعلى الرغم من قلة ما ذكرنا من دراسات عائد التكلفة في هذا الفصل ، فإن هذه الدراسات تمثل من المؤكد الاستثناء لا القاعدة ؛ ففي معظم المؤسسات قلما تبذل محاولات، وقد لا تبذل على الإطلاق ، للتعرف على عائدات أنشطة البحث في الإنتاج الفكري، على الرغم من بذل جهود من هذا القبيل في مجال الرعاية الصحية تفوق ما بذل في المجالات الأخرى . فقد تناول على سبيل المثال سكورا ودافيدوف Scura and Davidoff (1981) وجرينبرج وزملاؤه Greenberg et al. (1978) وشنول وولسون Schnall and Wilson (1976) الخدمات التي يقدمها المكتبيون العاملون في الطب الإكلينيكي ، حيث سألوا الأطباء الإكلينكيين المستفيدين من الخدمات عن مدى ما أحدثته المعلومات التي حصلوا عليها من أثر مباشر في رعاية المرضى . ومن خلال العمل في مكتبة أحد

المستشفيات طلب كنج (King 1987) من المستفيدين مما يقدم من خدمات ، الحكم على ما لهذه المعلومات من أهمية إكلينيكية ، وأهمية معرفية (إسهامها في تنمية معارف المستفيدين في مجال الرعاية الصحية) والحكم على نوعيتها ومدى حداثتها ، ومدى صلاحيتها بالنسبة للموقف الإكلينيكي الذي كان وراء طلب المعلومات ، وما لها من أثر في اتخاذ القرارات الإكلينيكية. ومما لا شك فيه بالطبع ، أن المرء قد يتطلع للذهاب إلى ما هو أبعد من ذلك، للتعرف على مدى ما يمكن أن تسهم به خدمة المعلومات الإكلينيكية في الحد من انتشار الأوبئة أو توفير العلاج المناسب ، أو الحد من الوقت الذي يمكنه المريض بالمستشفى ، أو الاقتصاد في تكاليف الرعاية الطبية.

ويقدم ولسون وستار - شنايكراوت وكوبر Wilson, Starr-Schneikraut and Cooper (1989) نتائج دراسة استخدم فيها أسلوب الواقعة الحاسمة، لتقييم عائدات عمليات البحث التي يتم إجراؤها في مرصد بيانات المدلاين ، حيث أجريت مقابلات هاتفية مع أكثر من ٥٠٠ من العاملين في مجال الصحة، ممن عرف عنهم الإفادة من المدلاين . وقد طلب ممن أجريت معهم المقابلات التركيز على عملية بحث حديثة في المدلاين . وفضلاً عن محاولة التعرف على ما كان لعمليات البحث الحديثة هذه من أثر في اتخاذ القرارات الطبية، حاول الباحثون التعرف على ما كان لها من آثار طويلة المدى. وانتهى الباحثون إلى أن « المعلومات التي أمكن الحصول عليها عن طريق المدلاين ، كانت لها منافعها الجمة التي أنقذت حياة بعض المرضى » . وقد استطاع الباحثون توثيق ثمانٍ من هذه الحالات الأخيرة.

ومن الصعب ، ولا شك ، إجراء تحليلات عائد التكلفة في بيئة خدمات المعلومات. وقد لا يكون لأي من الدراسات التي تنتمي إلى هذا النوع ، من المصادقية والثقة نصيب. إلا أنه ، أيّما كان الأمر ، فإنه يتعين على مرافق المعلومات العمل بشكل أو بآخر على تبرير وجودها . ومن ثم ، فإنه لا يمكن في الدراسات التقييمية ، تجاهل ما لخدمات هذه المرافق من عائدات ، حتى وإن بدت هذه العائدات ضبابية أو غير واضحة المعالم.

الفصل العاشر

العوامل المؤثرة على الأداء في استرجاع المعلومات

سبق أن ألمحنا إلى العوامل المؤثرة في كفاءة نظم استرجاع المعلومات ، كما تعرضنا لهذه العوامل بإيجاز ، في ثنايا هذا الكتاب . ويحاول هذا الفصل جمع شتات كل ذلك معاً ليقدّم صورة تتضمن المزيد من التفاصيل لما يؤدي للارتفاع بمستوى كفاءة نظام الاسترجاع أو تدني مستوى الكفاءة.

ويصور الشكل رقم (١٦) (في الفصل الثالث) تتابع الأحداث التي تقع بدءاً بتقدم المستفيد الذي يحتاج إلى معلومات في موضوع معين إلى مركز المعلومات ، إلى أن تقدم لهذا المستفيد مجموعة النتائج التي يسفر عنها البحث ^(١) . فينبغي أولاً ترجمة حاجة المستفيد إلى المعلومات ، إلى حاجة معرب عنها (بيان طلب) . ثم يتعين بعد ذلك على اختصاصي المعلومات اختيار أنسب مرصد البيانات التي يمكن إجراء البحث فيها ، وصياغة استراتيجية البحث المناسبة . ويتم إجراء عملية البحث بمضاهاة الاستراتيجية مقابل مرصد البيانات ، ويمكن لمسئول البحث تنقية النتائج للتخلص من المواد غير الصالحة ، قبل تقديمها للمستفيد.

وتنطوي كل خطوة من الخطوات التي يصورها الشكل رقم (١٦) على مصادر محتملة للوضوء أو الفاقد في المعلومات . هذا بالإضافة إلى الطابع التراكمي لأثر هذه

(١) البحث المفروض هو المفترض هنا . ويختلف الموقف حينما يقوم المستفيدون أنفسهم بإجراء عمليات البحث الخاصة بهم ، نظراً للقضاء على احتمالات الفاقد في المعلومات الذي يمكن أن ينتج عن إدراك الوسيط لحاجة المستفيد إلى المعلومات ، وترجمة هذه الحاجة أو الطلب إلى استراتيجية بحث ، أو ينتج عن تفويض عملية غربلة المخرجات . إلا أنه من الممكن للمستفيدين ألا يترجموا طلباتهم إلى استراتيجية بحث بدقة . هذا بالإضافة إلى أن الغالبية العظمى منهم قد لا يكونون متألفين مع إجراءات التعامل مع النظام وسياسات الكشف ومنطق البحث.

العوامل؛ فمن الممكن أن يكون هناك فاقد في المعلومات نتيجة لقصور الطلب عن التعبير بدقة عن حاجة المستفيد الحقيقية إلى المعلومات ، كما يمكن أن يكون هناك مزيد من الفاقد لأن المرصد الذي وقع عليه الاختيار ليس أنسب المراصد لإجراء البحث ، كذلك يمكن أن يكون هناك مزيد من الفاقد نتيجة لعدم ملاءمة استراتيجية البحث، وهكذا. وعند تحليل نتائج برنامج ما للتقييم ، وخاصة تحليل أخطاء الاستدعاء والتحقيق ، يحاول مسئول التقييم تحديد مواطن معظم المشكلات أو الأخطاء ، في دورة خدمة المعلومات.

وفي هذا الإطار يؤثر العديد من العوامل المهمة في كفاءة نظم استرجاع المعلومات الشكل رقم (٥٢). فهذه العوامل تؤثر في كفاءة (فعالية) عملية البحث ، كما تؤثر في التكاليف التي تتحملها المؤسسة أو يتحملها المستفيد ، ممثلة فيما ينفق من مال ووقت (قضيتا فعالية التكلفة وعائد التكلفة) .

وكانت العوامل الخاصة بالنظام، أو كما عبرت عنها سبارك - جونز Sparck-Jones (1981) « بخصائص بيانات الكشف وآليات البحث المتاحة » (P. 216) ، في إطار التطور التاريخي، هي التي تحظى بالدراسة من جانب الباحثين في مجال استرجاع المعلومات. إلا أن العوامل البشرية قد بدأت مؤخراً تخضع للتدقيق على نطاق واسع وبشكل مكثف. وتشمل هذه العوامل البشرية خصائص المكشفين ومسئولي عمليات البحث (كالخبرة في الكشف أو البحث ، والتخصص الموضوعي ، على سبيل المثال) وكلها يمكن أن تؤثر في كفاءة الاسترجاع . كذلك تطورت زوايا النظر في مقاييس التقييم المناسبة لكل دراسة بعينها. وقد تركز الاهتمام ، تقليدياً ، على مقاييس المخرجات كالاستدعاء والتحقيق والجدة والتكلفة (في المال والوقت) . وهناك الآن تزايد مستمر في اتجاه المهتمين بالتقييم نحو عملية البحث لا ناتج البحث . وقد استخدمت ، في حالات الاهتمام بدراسة عملية البحث، مقاييس تقييم أخرى كخصائص « النموذج السلبي » و « النموذج الإيجابي » التي تنسب لبورجمان (1984) Borgman ، حيث يتم تقييم قدرة مسئول البحث على تكوين نموذج عقلي ملائم لنظام البحث . وسوف نتناول هذه المقاييس بمزيد من التفصيل فيما بعد ، في هذا الفصل .

| عوامل النظام | العوامل البشرية |
|---|--|
| مرصد البيانات التغطية الفاصل الزمني وتتابع التجديد سياسات التكشيف ولغة التكشيف لغة التكشيف التخصيص مستوى الربط مدى الجهد الظاهر في البناء التكشيف الشمول التخصيص في المصطلحات المستعملة الدقة استراتيجية البحث الشمول التخصيص | التكشيف الإطار التمكن من الموضوع الخبرة في التكشيف البحث الاطراد التمكن من الموضوع الخبرة في البحث الأسلوب المعرفي التنقية الاطراد التمكن من الموضوع الخبرة في الغريلة |

الشكل رقم (٥٢) : العوامل التي تدرس في تقييم نظم استرجاع المعلومات.

العوامل الخاصة بالنظام في استرجاع المعلومات :

مرصد البيانات :

يتأثر اختيار مرصد البيانات (من جانب المستفيد النهائي أو الوسيط) أو اقتناء مرصد البيانات (من جانب متعهد مرصد البيانات) وبقوة بالمعايير التالية :

- المجال الموضوعي
- عوامل التكلفة
- الاعتبارات النوعية
- التغطية
- عوامل الوقت
- عوامل التكشيف واللغة

وبالنسبة للمستفيد من مرصدا البيانات ، تشمل عوامل التكلفة المهمة تكلفة الوحدة لكل تسجيلية مسترجعة ، وتكلفة الوحدة لكل تسجيلية صالحة مسترجعة. أما بالنسبة للمؤسسات التي تقوم بإنشاء مرصدا البيانات وتسويقها ، فإن أبرز عوامل التكلفة هي:

- تكلفة إنشاء مرصد البيانات
 - تكلفة الحصول على مرصد البيانات
 - تكلفة تركيب مرصد البيانات (تكاليف التركيب والتشغيل)
 - تكلفة الوحدة لكل تسجيلية ، والتكلفة بالنسبة لعدد المداخل أو نقاط الوصول المتوافرة
 - التكلفة بالنسبة لحجم الطلب على الخدمة (تكلفة الوحدة لكل عملية بحث)
 - التكلفة بالنسبة للخصائص النوعية (كالتغطية ومدى الشمول في الكشف مثلاً)
- وللمعايير النوعية أهميتها نظراً لأنها تؤثر في كل من نوعية نتائج عملية البحث ومدى اكتمال هذه النتائج . والأنواع الثلاثة الرئيسية للمعايير النوعية هي التغطية وعامل الوقت ، وعوامل الكشف واللغة (الشكل رقم (٥٣)).

التغطية :

تشمل معايير التغطية التي مهدنا لها في الفصل السابع ، مدى اكتمال مرصد البيانات، والاكتمال النسبي لاثنتين أو أكثر من مرصدا البيانات ، ومدى التداخل أو التكامل بين اثنتين أو أكثر من مرصدا البيانات . ومن الممكن النظر في الاكتمال بناء على العدد النهائي للوثائق التي يتم كشفها أو استخلاصها سنوياً ، وعدد المصادر (الدوريات على سبيل المثال) التي تحظى بالتغطية ، ونوعية المصادر (كالتقارير التقنية وبراءات الاختراع مثلاً) والفترة الزمنية المغطاة . ولا يرتبط هذا الاعتبار الأخير إلا بأهمية مرصد البيانات بالنسبة للبحث الراجع. فلا قيمة تذكر لمرصد البيانات بالنسبة للبحث الراجع، إلا إذا كان يغطي الإنتاج الفكري لثلاث سنوات على الأقل. وربما يبدأ في الاقتراب من المستوى المرتفع للأهمية عندما يصل إلى تغطية خمس سنوات على وجه التقريب. ويمكن لفترات التغطية الأطول من ذلك أن تكون أكثر أهمية في العلوم الاجتماعية والإنسانيات منها في العلوم والتقانة، نظراً لأن معدل التعطل عادة ما يكون أعلى في المجالات العلمية والتقنية مما هو عليه في المجالات الأخرى.

| عوامل التغطية | عوامل الوقت | عوامل الكشف واللغة |
|--|--------------------------------|--|
| عدد المصادر نوعية المصادر عدد المواد المدة الزمنية مدى الاكتمال بالنسبة لاحتياجات المستفيدين التفرد والتداخل | الفاصل الزمني تتابع التجديد | درجة التحكم في اللغة مدى التخصيص في اللغة أدوات البحث المتوافرة مدى الغموض الدلالي والنظمي مدى الشمول (عدد نقاط الوصول وأنواعها) الدقة والاطراد (الأخطاء الملاحظة) |

الشكل رقم (٥٣) : بعض الجوانب النوعية الأساسية لمرصد البيانات.

ومن الصعب بمكان للأسف قياس التغطية المطلقة لمرصد البيانات في مجال تخصصي معين، وذلك بمجرد إحصاء عدد الوثائق أو عدد المصادر التي يتم تغطيتها. ويتطلب تقدير مدى التغطية تقييماً محكماً سبق أن عرضنا له في الفصل السابع. وكان الأسلوبان اللذان أمكن التحقق منهما:

(١) طريقة « الوراقية » ، حيث يتم تجميع عينة عشوائية من المقالات الصالحة، حيث تتمتع المراجعات العلمية بأهمية خاصة لهذا الغرض، ثم اتخاذ الإشارات الوراقية الواردة في هذه العينة من الوثائق ، « كتجمع للاستشهادات المرجعية » لاختبار تغطية مرصد بيانات معين.

(٢) طريقة « السمات » ، حيث يتم تقييم التغطية بالبحث عن الوثائق المتصلة بالاستفسارات الموضوعية في عدد من مرصد البيانات.

وقد تبين لدافيسون وماثيوز (Davison and Matthews (1969) اعتماداً على وراقية تشتمل على ١٨٢ تسجيلية حول موضوع الحاسبات الآلية في قياس أطياف الكتل، أنه من بين اثني عشر مرصداً للبيانات لم يغطِ مرصد واحد أكثر من ٤٠ بالمائة من هذا الإنتاج الفكري، وأن تغطية الـ *Chemical Abstracts* كانت ٢٤ بالمائة فقط. ووفقاً لتقدير بورن (Bourne (1969 a تبين أن الـ *Bibliography of Agriculture* لا تغطي سوى ٥٠-٦٠ بالمائة من الإنتاج الفكري المتصل بالبحوث الزراعية. واعتماداً على

تجمع للاستشهادات المرجعية المنتقاة من الدوريات المتخصصة في الطب الشرعي، حاول يونكر وآخرون (1990) Younker et al. مؤخراً، التعرف على تغطية كل من MEDLINE و EXCERPTA MEDICA و SCISEARCH و CRIMINAL JUSTICE و PERIODICAL INDEX، وما بينها من تداخل. وقد تبين لهم أن الـ EXCERPTA MEDICA قد حققت أفضل تغطية (٩١ بالمئة) وأن CRIMINAL JUSTICE PERIODICAL INDEX يسجل أسوأ تغطية (٢٤ بالمئة) . وقد أمكن العثور على جميع الاستشهادات المرجعية في واحد على الأقل من مرصد البيانات.

واستخدمت تينوبير (1982) Tenopir مجموعة « سمات » عامة جداً قصد بها أن تكون بحثاً شاملاً حقيقياً عن موضوع إدارة الطوارئ العريض. وفي مقابل ذلك اتبع كل من شارما (1982) Sharma ونيكسون (1989) Nixon طريقة مختلفة، حيث اختارا مجموعة من الاستفسارات المحددة في مجالي تنمية المجتمع والتغذية البشرية، على التوالي. وقد تبين للباحثين الثلاثة أنه يتعين اللجوء إلى أكثر من مرصد بيانات واحد لضمان شمول عملية البحث. كذلك قارنت تينوبير طريقتهما الخاصة بمجموعة السمات العامة بطريقة الوراقية، وتبين لها أن كلا الطريقتين واحد تقريباً، من حيث تقديرات التغطية.

ولدراسات التغطية أهميتها البالغة نظراً لأنها تدل على أنه ليس هناك مصدر واحد شامل تماماً في تغطيته للإنتاج الفكري في مجال بعينه. ويؤكد جهد برادفورد (1953) ford وما تلاه من تحليلات قياسوراقية أنه من الممكن لكثير من المصادر الأولية أن تسهم في الإنتاج الفكري لمجال موضوعي معين، على الرغم من أنه يمكن لنسبة مئوية عالية من هذا الإنتاج الفكري أن ترد في عدد محدود جداً من المصادر.

وهناك بعض جوانب أخرى للتغطية جديرة بالذكر؛ فمن المهم التعرف على مدى إمكان الاعتماد على خدمة المعلومات في تغطية الإنتاج الفكري. ومما لا شك فيه أنه يتعين على هذه الخدمة التدقيق بانتظام في جميع أعداد الدوريات التي تدعي تغطيتها، كما ينبغي أن تميز بين الدوريات التي تغطيها تغطية كاملة وتلك التي تغطيها انتقائياً. وحيثما يتم الاختيار من مصادر متعددة، فإنه من المهم أيضاً تقييم مدى اطراد عملية الاختيار ومدى إمكانية الاعتماد عليها. ولا يمكن تحقيق هذا النوع من التقييم أيضاً بمجرد الملاحظة وحدها، وإنما عن طريق الاختبار.

وهناك اعتبار آخر في اختيار مرصد البيانات يتصل بتوافر المصادر المغطاة محلياً. ومن الممكن مقارنة قائمة المصادر التي تحظى بالتغطية في مرصد بيانات معين، بالمصادر المتاحة محلياً. وقد قام كل من بوشانان و برونند وكارلين Buchanan, Ber- (1989) wind and Carlin بمقارنة المصادر المسجلة على أسطوانات ضوئية مكتنزة، وذلك لا للتعرف على عدد ما تقتنيه المكتبة من الوثائق المسترجعة فحسب، وإنما للتعرف أيضاً على عدد ما كان بإمكان المستفيدين من المكتبة التحقق منه والعثور عليه. وتؤدي الإفادة من خدمات الإحاطة الجارية والبحث الراجع في مجتمع معين إلى نشأة الحاجة إلى نظام للإمداد بالوثائق، ومن المهم بمكان تقييم أثر هذه الخدمات على غيرها مما تقدمه المؤسسة من خدمات، كتبادل الإعارة بين المكتبات مثلاً.

عوامل الوقت :

مما لا شك فيه أنه من المهم بمكان لكثير من المستفيدين من مرصد البيانات، أن يكون المرصد حديثاً. ويؤثر الفاصل الزمني بين نشر الوثيقة وظهورها في مرصد البيانات، على أهمية المرصد لأغراض الإحاطة الجارية. فينبغي لخدمة البث الانتقائي للمعلومات، الجيدة أن تعمل بمعدل جده مرتفع؛ أي أنه لا ينبغي للغالبية العظمى مما يقدم للمستفيدين من وثائق أن يكون صالحاً لاهتماماتهم فحسب، وإنما ينبغي أيضاً أن يكون من الوثائق التي لم يسبق لهم الاطلاع عليها. ومن الواضح تأثر معدل الجدة (عدد الوثائق الصالحة المسترجعة الجديدة بالنسبة للمستفيد / عدد الوثائق الصالحة المسترجعة) وإلى حد بعيد بمدى سرعة تجهيز الوثائق في مرصد البيانات.

ومرصد البيانات التي يتضاءل فيها التجهيز الفكري بواسطة البشر، بوجه عام، إلى أدنى حد، كما هو الحال مثلاً في تلك التي ينتجها معهد المعلومات العلمية Institute for Scientific Information، أكثر حداثة وتجديداً من تلك التي تتطلب المزيد من التجهيز الفكري ككتابة المستخلصات مثلاً. وقد تبين لأشمول و سميث وشترن Ashmole, (1973) Smith and Stern عند تقييم مرصد البيانات وفقاً لأهميتها بالنسبة للصناعات الصيدلية، أن مرصد بيانات ASCA الذي ينتجه معهد المعلومات العلمية، كان يشتمل على الوثائق بعدما يتراوح بين أقل من الأسبوع وثلاثة أسابيع من النشر، بينما تراوحت الغالبية العظمى من الخدمات الأخرى بين الشهرين والستة أشهر، في حين تراوحت الـ Biological Abstracts بين أربعة أشهر واثنى عشر شهراً.

ويتصل بمدى الحداثة التتابع الذي يتم به تجديد مرصد البيانات ، فإذا كان المرصد مرتبطاً بدورة للنشر ، فإن تتابع التجديد يمكن أن يكون متوافقاً مع تتابع النشر ، سواء كان نصف شهري ، أو شهرياً ... إلخ . وتختلف تكلفة الإيجار بالنسبة لبعض الخدمات، تبعاً لمدى حاجة المستفيد إلى سرعة تجدد مرصد البيانات.

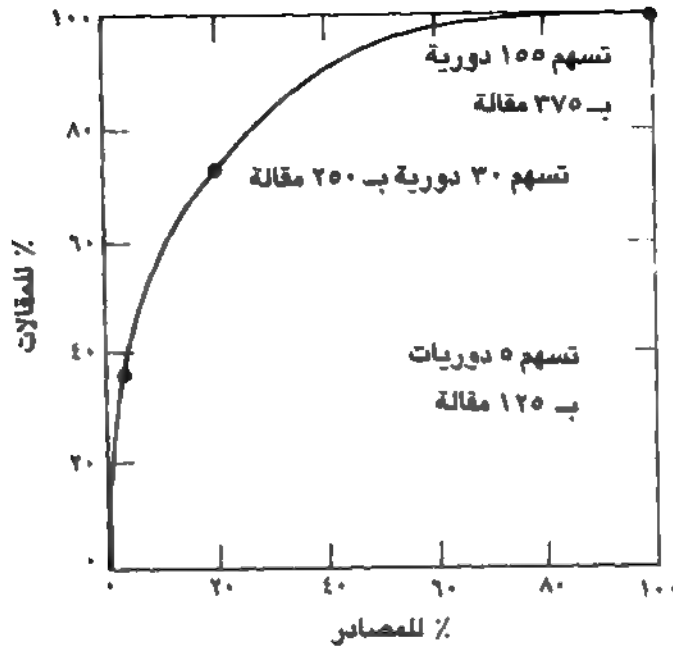
التكشيف واللغة :

تحتاج بعض العوامل الخاصة بالتكشيف واللغة ، الواردة في الشكل رقم (٥٣) إلى بعض التوضيح؛ فالتخصيص يتأثر بما إذا كانت اللغة المستعملة مقيدة أم لغة طبيعية . وللغة المقيدة أثرها المرغوب ، في تجميع الوثائق التي تدور حول الموضوع نفسه تحت المصطلح نفسه، كما يمكن أن تكفل أيضاً أدوات مساعدة لمستول البحث من خلال بنيتها الدلالية الظاهرة. وفي مقابل ذلك، فإن اللغة المقيدة ليست مخصصة وليست قادرة على مواكبة التطورات الجارية بالقدر نفسه الذي تتمتع به اللغة الطبيعية. والشمول ببساطة هو مدى الإحاطة بالموضوعات التي تشملها الوثيقة والتعبير عن هذه الموضوعات بمصطلحات كشفية. أما الدقة فتدل على نوعية التكشيف المطبق في مرصد البيانات.

وقد أجري الكثير من الدراسات حول أثر لغة التكشيف وسياسات التكشيف على كفاءة الاسترجاع ؛ فقد درس باربر وآخرون (Barber et al. (1988) التكشيف في مراصد البيانات الصيدلية ، في حين درس ستار (Starr (1982) التكشيف في مراصد البيانات الخاصة بعلوم البحار. وقد انتهت هذه الدراسات إلى أنه طالما أمكن التعرف على سياسة التكشيف فإنه يمكن التنبؤ بمدى ملائمة كل مرصد من مراصد البيانات بالنسبة لنوعيات بعينها من عمليات البحث، نظراً للتنوع الكبير في سبل معاملة الموضوعات في مختلف مراصد البيانات . وبتركيزه على الكلمات المفردة في تسجيلات علم الحاسب الآلي، تبين لكوننجز (Konings (1985 كيف استعملت كلمات مختلفة تمام الاختلاف للتعبير عن المفاهيم أو الموضوعات نفسها في مختلف مراصد البيانات .

الجوانب الخاصة بفعالية تكلفة مرصد البيانات:

يحتاج متعهد مراصد البيانات إلى التعرف على الحد الأقصى للإنتاج الفكري الذي يقوم بتجميعه وإتاحته في مجال موضوعي معين. ويتعلق قانون برادفورد للتشتت،



الشكل رقم (٥٤) : توزيع برادفورد لـ ٣٧٥ مقالة صدرت في ١٥٥ دورية .

ولتحليل برادفورد أهميته البالغة في بيان كيفية توزيع الموارد المالية لمركز المعلومات بأقصى درجات الكفاءة في اقتناء أوعية المعلومات. ويوحى المنحنى الوارد في الشكل رقم (٥٤) باستراتيجيتين محتملتين للاقتناء. فإذا افترضنا أن هناك من يقوم بإنشاء مركز للمعلومات في مجال فرط الموصلية، وكانت هناك ميزانية قدرها ص من الدولارات متاحة لاقتناء الدوريات، فإن أسلم استراتيجية هي شراء الدوريات وفقاً لإنتاجها المتوقع من المقالات الصالحة إلى أن تستنفد الميزانية. ففي المثال الوارد في الشكل رقم (٥٤) من المتوقع أن تنتج ٣٠ دورية ثلثي الإنتاج الفكري المناسب، بينما يمكن أن يتطلب الأمر أكثر من ١٠٠ دورية إضافية لتحقيق التغطية بنسبة ١٠٠ بالمئة. أضف إلى ذلك أن مجموعة الدوريات عالية الإنتاجية يمكن أن تكون مستقرة نسبياً، لفترة زمنية قصيرة على الأقل، أما قائمة الدوريات منخفضة الإنتاجية فيمكن أن تكون بعيدة كل البعد عن الاستقرار؛ فقائمة الدوريات التي تسهم كل منها بمقالة واحدة فقط في موضوع فرط الموصلية عام ١٩٩٢ م، يمكن أن تكون مختلفة تمام الاختلاف عن قائمة تلك الدوريات التي تسهم كل منها بمقالة واحدة فقط عام ١٩٩٣ م.

ولا يمكن بالنسبة لمعظم مراكز المعلومات في الأساس تحقيق التغطية بنسبة ١٠٠ بالمئة، وحتى إن أمكن تحقيقها، فإن ذلك يمكن أن يكون هدفاً غير معقول من وجهة نظر فعالية التكلفة. وللاارتفاع بالتغطية من ٨٥ بالمئة مثلاً إلى ما يقترب من ١٠٠ بالمئة،

فإن الأمر قد يتطلب مستوى إنفاق غير متناسب على الإطلاق، نظراً لأن الـ ١٥ بالمائة الأخيرة للتغطية يمكن أن تبلغ تكلفتها ببساطة قيمة تكلفة الـ ٨٥ بالمائة الأولى، أو أكثر. واستراتيجية الشراء الأخرى إذن هي وضع هدف واقعي إلى حد ما للتغطية، ربما يصل إلى ٨٥ أو ٩٠ بالمائة، ثم اتباع توزيع برادفورد الذي يمثل قانوناً لتناقص العائد، وذلك لاختيار تلك الدوريات التي تتمتع بأعلى احتمالات تحقيق الهدف.

وقد تناول عدد من الدراسات اتباع قاعدة ٨٠ / ٢٠ في قرارات الاقتناء. وهذه تشمل الدراسات الخاصة ببيانات الإعارة لاختيار الكتب (Britten, 1990) وتحليل إحصاءات الاطلاع الداخلي على الدوريات لاتخاذ قرارات الاشتراك وإلغاء الاشتراكات (Alligood, Russo-Martin and Peterson (1983) .

ولتوزيع برادفورد عندما يقترن « بقانون التعطل » أهميته أيضاً في بيان التوزيع الأمثل لحيز الاختزان المتاح ، بحيث تكون الوثائق التي يمكن أن يستفاد منها أكثر هي الأقرب إلى متناول المستفيدين. وتكمن المشكلة كما هو واضح في تحديد أي الأوعية يمكن أن يشتد عليها الطلب وأنها يمكن أن تكون العكس.

وينص قانون التعطل على أن احتمال الطلب على أوعية المعلومات يتناقص تبعاً لعمرها، وخاصة في العلوم والتقانة . فقد تبين لفوسلر وسايمون Fussler and Simon (1969) أثناء عملهما بجامعة شيكاغو، أنه من الممكن للكتب أن تحوّل، دون أدنى مخاطرة، إلى أماكن التخزين بناء على اللغة وتاريخ النشر فقط. وغالباً ما يتم التعبير عن معدل تعطل أوعية المعلومات أو تقدمها في السن، بناء على « منتصف الإفادة » منها . وعمر نصف الإفادة من قطاع معين من الإنتاج الفكري هو عدد السنوات اللازمة بشكل راجع لتلبية نصف إجمالي الطلب على الإنتاج الفكري في الموضوع ، أو لاجتذاب نصف إجمالي الاستشهادات المرجعية بالإنتاج الفكري في السنة الجارية (وسط عمر الاستشهاد المرجعي) . وعلى ذلك فإنه إذا كان نصف إجمالي الاستشهادات المرجعية الواردة في الإنتاج الفكري لعام ١٩٩٣ م في موضوع المواد فائقة الموصلية تشير إلى الإنتاج الفكري الصادر خلال الاثنين والأربعين شهراً الماضية، فإن وسط عمر الاستشهاد المرجعي لهذا الإنتاج الفكري يبلغ ثلاثة أعوام ونصف العام. وكذلك الحال أيضاً، فإذا كان ٥٠ بالمائة من إجمالي الطلب على الإنتاج الفكري للمواد فائقة الموصلية، في إحدى المكتبات المتخصصة في الفيزياء، يتعلق بالإنتاج الفكري الصادر خلال الأشهر

الاثنين والأربعين الأخيرة ، فإن وسط عمر الإفادة من هذا الإنتاج الفكري يبلغ ثلاثة أعوام ونصف العام. ومن المفترض في غالب الأحيان أن تسفر هاتان الطريقتان لقياس التعطل عن نتائج متناظرة تقريباً ، وبعبارة أخرى ، فإنه يمكن الاعتماد على أنماط الاستشهاد المرجعي للتنبؤ بأنماط الطلب على أوعية المعلومات المكتبات (Bourne, 1965; Fussler and Simon, 1969) وإن كان هناك من الباحثين من يشكك في ذلك التأكيد (Sandison, 1971; Line and Sandison, 1974)

ونعود إلى مشكلة توزيع الحيز ، ونفترض أن لدى مركز معلومات معين حيزاً على أرففه المفتوحة يكفي لاستيعاب ٣٠٠٠ مجلد من الدوريات. وينبغي توزيع هذا الحيز بحيث تكون المجلدات المختزنة به هي التي يشتد الطلب عليها من جانب المستفيدين. ومن الممكن الربط بين دراسات عوامل تقدم السن ودراسات التوزيع المتوقع للطلب على عناوين بعينها، للتعرف على أفضل سبل توزيع هذا الحيز. فالدورية ص على سبيل المثال يمكن الاحتفاظ بها على الأرفف المفتوحة ولمدة عشر سنوات راجعة، والدورية ع يمكن الاحتفاظ بها في المكان نفسه ولكن لمدة ثلاث سنوات راجعة فقط. ويمكن لمثل هذه التحليلات العملية لفعالية التكلفة أن تدل على أن تلك المجلدات التي يتم الاحتفاظ بها في أقرب مكان في متناول المستفيد ، هي تلك التي يتوقع لها أن تلبي نسبة مئوية معينة من طلبات المستفيدين. وبإمكان منتجي مرصود البيانات الاعتماد على هذا النوع من الدراسات لتحديد المجموعة المثلى من الوثائق التي يتم تجميعها وتكثيفها ، فضلاً عن تحديد المدى الزمني الراجع لإتاحة مرصود البيانات في أيسر أشكاله مثلاً، كتحديد المدى الزمني الراجع لإتاحة المرصود على الخط المباشر مثلاً.

وينبغي إجراء مثل هذه الدراسات في مرحلة التصميم وقبل البدء في تنفيذ النظام. وبعبارة أخرى فإن الأمر يتطلب التنبؤ بفعالية التكلفة الخاصة بمختلف فئات أوعية المعلومات. وهذه المهمة أصعب من تقييم عوامل الإفادة في نظام عامل فعلاً. إلا أن هناك عدداً من الأساليب التي تكفل إمكانية هذا النوع من التنبؤ، ومن بينها إحصاء الاستشهادات المرجعية ، وتحليل الإفادة من المكتبة أو تحليل حركة تبادل الإعارة بين المكتبات، وتحليل الاستفسارات الموضوعية التي تتجمع من عينة من مجتمعات المستفيدين. ويمكن لكل هذه التحليلات أن تساعد في التنبؤ باحتمال الإفادة من مختلف أنواع أوعية المعلومات.

لغة التكشيف :

سبق أن تناولنا في الفصل السادس الخصائص العامة للغة المستعملة في الاسترجاع، وهذه تشمل :

١ - مدى التخصيص في اللغة ؛ ما إذا كانت اللغة المستعملة في التكشيف والبحث تميل إلى التعميم أم التخصيص.

٢ - مستوى الربط في اللغة ، أي مدى الربط المسبق للكلمات المفردة في صيغ أكثر تعقداً، باستعمال العبارات أو الرؤوس الفرعية أو الأدوات اللغوية الصريحة كالروابط أو الأدوار.

٣ - مقدار ما يقدم صراحة للكشف أو مسئول البحث من تنظيم للعلاقات، في شكل علاقات دلالية بين المصطلحات (كالمصطلحات الأعرض أو الأضيّق أو المتصلة مثلاً) ويساعد هذا التنظيم البنائي في اختيار المصطلحات في التكشيف وفي صياغة استراتيجيات البحث.

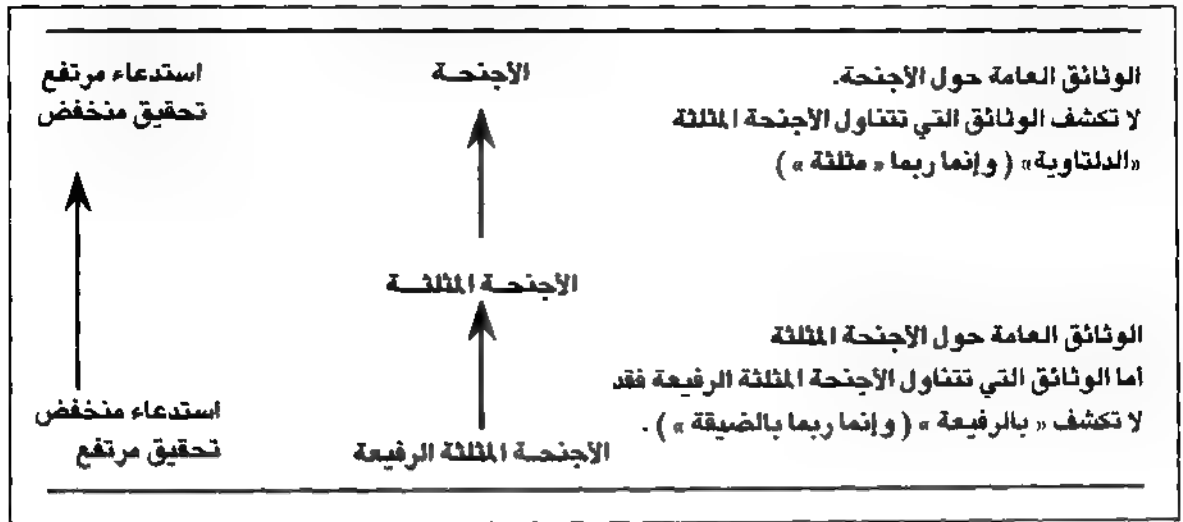
التخصيص :

التخصيص في اللغة هو أهم العوامل المؤثرة في مدى التحقيق الذي يمكن أن يتم به إجراء البحث على الإطلاق. فإذا افترضنا أن أحد المستفيدين يبحث عن المعلومات حول البطاقات متناهية الصغر، وكان النظام يشتمل على المصطلح المخصص ULTRA-MICROFICHE ، فإنه من الممكن إجراء بحث على مستوى عالٍ من التحقيق ، حيث يمكن لمعظم إن لم يكن كل الوثائق المسترجعة أن تكون متصلة بالموضوع المخصص للاستفسار. ولكن النظام قد لا يشتمل إلا على المصطلح الأعم MICROFICHE فقط، ومن ثم فإنه من المحتم أن يسفر البحث تحت هذا المصطلح، بينما يطلب المستفيد البطاقات متناهية الصغر على وجه التخصيص، يسفر عن انخفاض في نسبة التحقيق، لأنه من الممكن لمعظم الوثائق المسترجعة ألا تكون متصلة بالموضوع المخصص للبحث. وربما يزداد الموقف سوءاً إذا كان أكثر المصطلحات المتاحة تخصيصاً هو مصطلح المصغرات الفيلمية MICROFORMS .

ولننظر في ثلاثة نظم مستقلة للاسترجاع أ و ب و جـ، تستعمل لغات مقيدة

تشتمل على ٢٠٠٠ و ١٠٠٠ و ٥٠٠ مصطلح كشفي على التوالي. ونفترض أن مجموعة الوثائق المتخصصة في الديناميكا الهوائية نفسها كشفت في كل من النظم الثلاثة. وبعض هذه الوثائق يتناول موضوع الأجنحة المثثة الرفيعة **Slender delta wings**.

عند تكشف مثل هذه الوثائق في النظام أ الذي يستعمل لغة تشتمل على ٢٠٠٠ مصطلح كشفي، فإنه قد يكون من الممكن التعبير عن محتواها الموضوعي بشكل دقيق بواسطة وسيمة الفئة **SLENDER DELTA WINGS**. وعند كشف الوثائق نفسها في النظام ب الذي يستعمل ١٠٠٠ مصطلح كشفي فقط، فإنه قد لا يكون من الممكن التعبير عن محتواها الموضوعي بدقة، وربما لا يكون هناك بديل لتصنيفها تحت وسيمة الفئة الأعم، الأجنحة المثثة **DELTA WINGS**. أما في النظام جـ فإن الأجنحة **WINGS** قد يكون هو أكثر المصطلحات المتوافرة تخصيصاً لتكشف هذه الفئة من الوثائق. ويوضح الشكل رقم (٥٥) ما لهذا التفاوت في التخصيص من أثر على الاسترجاع.



الشكل رقم (٥٥) : اثر التخصيص في اللغة على كفاءة نظام الاسترجاع.

وعند سؤال النظام أ عن الوثائق المتصلة بموضوع الأجنحة المثثة الرفيعة، فإنه يمكن التعبير عن الموضوع في استراتيجية البحث بنفس دقة التعبير عنه في الكشف، ويمكن أن نتوقع تمتع مجموعة الوثائق الفرعية المسترجعة بنسبة تحقيق مرتفعة، إلا أنه من المحتمل أن تكون نسبة الاستدعاء منخفضة، لأن النظام يمكن أن يحجب عدداً من الوثائق التي تشتمل على معلومات مفيدة في موضوع الأجنحة المثثة الرفيعة. وبالسؤال عن فئة الوثائق التي تتناول الأجنحة المثثة الرفيعة على وجه التخصيص، فإن مسئول البحث يفوته استرجاع :

١- الوثائق التي تتناول الأجنحة المثلثة بوجه عام ، ولكنها تشتمل على قدر كبير من المعلومات حول الأجنحة المثلثة الرفيعة (والتي كشفت تحت الأجنحة المثلثة) .

٢ - الوثائق التي تتناول الأجنحة المثلثة الرفيعة ، ولكنها كشفت تحت بعض المصطلحات أو تجمعات المصطلحات المؤلفة، المتصلة دلاليًا بالموضوع كالأجنحة المثلثة الضيقة مثلاً .

وعند إدخال الطلب نفسه في النظام ب فإن الاستدعاء يميل للتحسن ؛ حيث يتم استرجاع بعض الوثائق الصالحة الإضافية التي أخطأها البحث في النظام أ ، إلا أنه من الممكن لبعض الوثائق التي يحتمل أن تكون مفيدة ، كتلك التي تتناول مثلاً الأجنحة بوجه عام وتشتمل على معلومات عن الأجنحة المثلثة الرفيعة ، وتلك التي تتناول الأجنحة المثلثة والتي كشفت تحت بعض المصطلحات أو تجمعات المصطلحات المؤلفة، المتصلة دلاليًا بالموضوع، كالأجنحة الدلتاوية مثلاً ، من الممكن لهذه الوثائق أن يخطئها البحث. وربما أمكن استرجاع هذه الوثائق الإضافية المفيدة بالبحث في النظام ج ، والذي يُضطر فيه الافتقار إلى التخصيص في لغة التكشيف، السائل للبحث تحت الفئة العامة العريضة « أجنحة » .

ومع انتقال مسئول البحث من النظام أ إلى النظام ب إلى النظام ج ، يتناقص عدد وسيمات الفئات المتميزة واضحة المعالم، ويزداد حجم فئات الوثائق، ولهذا يترادف عدد الوثائق التي يتم استرجاعها في كل عملية بحث، ويتحسن أداء الاستدعاء ، ويميل أداء التحقيق في الوقت نفسه للانخفاض. فاستجابة لطلب الوثائق التي تتناول الأجنحة المثلثة الرفيعة، ينبغي أن تشتمل الفئة الموسومة « بالأجنحة المثلثة الرفيعة » على وجه التحديد، على نسبة من الوثائق المفيدة أعلى من تلك التي تشتمل عليها الفئة الموسومة «بالأجنحة المثلثة» والتي تشتمل بدورها على نسبة من الوثائق المفيدة أعلى من تلك التي تشتمل عليها الفئة الموسومة « بالأجنحة » .

وكما ازدادت لغة التكشيف تخصيصاً ازدادت إمكانية التعبير عن الموضوع بدقة، وتناقصت أحجام فئات الوثائق التي تنشأ على هذا النحو ، ومن ثم ترتفع نسبة التحقيق التي يمكن بلوغها بالبحث في النظام. إلا أن الاستدعاء يكون منخفضاً لأن فئات الوثائق التي يحددها النظام عادة ما تكون صغيرة الحجم .

وترتبط مشكلة الافتقار إلى التخصيص بنظم اللغات المقيدة بوجه خاص؛ نظراً لأن اللغة المقيدة تعني بطبيعتها الافتقار إلى التخصيص الكامل؛ فمثل هذه اللغة ليست بالمجموعة الكاملة للمصطلحات المتوافرة في المجال الموضوعي، وإنما مجرد مجموعة فرعية محدودة من هذه المصطلحات لا أكثر. فقائمة رءوس الموضوعات الطبية *Med-ical Subject Headings (Me SH)* المستخدمة من جانب المكتبة القومية للطب، على سبيل المثال، تشتمل على عدة آلاف من المصطلحات، إلا أنها أصغر بكثير جداً من مفردات أي معجم طبي معياري. وفي مقابل ذلك، فإن المحتوى الموضوعي للوثيقة، في النظام المعتمد على اللغة الطبيعية، يتم التعبير عنه بمستوى التخصيص نفسه الذي تتمتع به الكلمات الواردة في الوثيقة نفسها. ولهذا السبب، يميل البحث في اللغة الطبيعية لارتفاع بمستوى التحقيق، كما تأكد فعلاً في عدد من الدراسات، في مختلف المجالات الموضوعية، بما في ذلك الكيمياء (Rowlett, 1977) والقانون (Carrow and Nugent, 1977).

ومن المشكلات المهمة التي تواجه مديري أي نظام لاسترجاع المعلومات، تحديد مدى الحاجة إلى التخصيص في اللغة. فاللغة ينبغي أن تكون مخصصة بما يكفي لكفالة إجراء الغالبية العظمى من عمليات البحث على مستوى مقبول للتحقيق. وتعني هذه العبارة أن مستوى التخصيص يتفاوت من لغة لأخرى، حيث أن هناك بعض المجالات الموضوعية التي تنمو بشكل أكثر تفصيلاً من غيرها. فاللغة التي وضعتها المكتبة القومية للطب قد تحتاج فقط إلى عدد قليل من المصطلحات العامة في الرياضيات، واللغة التي تضعها الجمعية الرياضية الأمريكية قد تحتاج فقط إلى عدد قليل من المصطلحات العامة ذات الطبيعة الطبية. وتعني هذه العبارة أيضاً أن اللغة ينبغي أن تعدل باستمرار لكي تصبح أكثر تخصيصاً، استجابة لما ينشر من إنتاج فكري أكثر تخصيصاً، أو ما يمكن أن يقدم للنظام من طلبات أكثر تخصيصاً. وينطوي ذلك بدوره على معنى ضمني؛ فمديرو النظام ينبغي أن يتلقوا تلقياً مرتداً من المستفيدين يمكنهم من التعرف على ما يكتنف اللغة من قصور. وعلى ذلك، فإن الأمر يتطلب نوعاً من التقييم المستمر أو ضرباً من إجراءات ضبط الجودة، وهو ما لا يدخل ضمن المكونات الروتينية لكثير من خدمات المعلومات.

وتعديل مستوى التخصيص لإحدى اللغات القائمة أيسر من بلوغ المستوى الملائم من التخصيص في المراحل الأولى لإعداد المكنز. فمنذ عام ١٩١١م أدخل هالم Hulme مبدأ مهماً يعرف « بالمسوغ الأدبي » أو « المسوغ الوراقى » . فهذا المبدأ الذي طبقه هالم (1911 a, b, c) على تصنيف الكتب، ينص ببساطة على أن المصطلح يصبح مسوغاً أو له ما يبرره إذا كان من المعروف وجود إنتاج فكري في الموضوع الذي يدل عليه. ومن الممكن للمبدأ أن يمتد إلى إعداد المكنز، بالقول بأن المصطلح يصبح مسوغاً إذا تأكد وجود إنتاج فكري كافٍ في الموضوع الذي يدل عليه، وأنه ما لم يدخل المصطلح في اللغة يمكن للإنتاج الفكري أن يتوارى بين ثنايا فئة أكبر بكثير، وبذلك يصبح من الصعب استرجاعه.

وهناك مبدأ ملازم مهم ، مبدأ غالباً ما تغفل عنه ، يمكن أن يسمى « مسوغ المستفيد »؛ فالمصطلح يصبح له ما يبرره إذا كان من الممكن لطلبات المعلومات على هذا المستوى من التخصيص، أن تقدم وبدرجة عالية نسبياً من الكثافة ، من جانب المستفيدين من النظام. وربما كان مسوغ المستفيد أهم من المسوغ الوراقى في وضع اللغات المقيدة، لأغراض استرجاع المعلومات، على درجة عالية من الكفاءة. فربما كان من الممكن إعداد قائمة تضم عدة مئات، من فصائل الكلاب، يمكن أن يكون لكل منها ما يسوغه وراقياً، ولكن إذا كان المستفيدون من نظام معين لا يمكن أن يحتاجوا إلى ما هو أكثر تخصيصاً من « الكلاب » ، فلن يكون هناك مبرر مقبول لتطوير هذا الجزء من اللغة بهذا القدر الكبير من التفصيل. ومعنى ذلك واضح؛ فمصمم اللغة المقيدة ينبغي أن يتوافر له قدر كبير جداً من المعلومات حول المستفيدين المحتملين من النظام، وحول أنواع ما يمكن أن يتقدموا به من طلبات، وهذه نقطة يؤكد بها بقوة كل من سورجل Soergel (1974) ولانكستر (1972) Lancaster .

ويمكن لافتقار لغة التكشيف إلى التخصيص أن يكون سبباً في أخطاء الاستدعاء أو التحقيق. فإذا كانت هناك فئة معينة من الوثائق لم يتم التعبير عنها بشكل مخصص، وإنما استعمل مصطلح من مصطلحات الدخول إلى اللغة ليدل على كيفية تصنيف الفئة ضمن فئة أكبر، فإن أخطاء التحقيق سوف تحدث نظراً لافتقار اللغة إلى التخصيص. أما إذا كان التجاهل من نصيب الموضوع حتى في مصطلحات الدخول إلى اللغة، فإن ذلك

يؤدي إلى أخطاء في كل من الاستدعاء والتحقيق. لننظر في موضوع ظواهر اكتمال الإدراك الحسي، ونفترض أنه لا يمكن التعبير عن هذا الموضوع بشكل مخصص، وإنما يمكن تكشيفه تحت مصطلحين مؤتلفين هما «الرؤية والأوهام». وينبغي تسجيل هذا القرار في مصطلحات الدخول إلى اللغة هكذا :

PERCEPTUAL COMPLETION PHENOMENA

use VISION and ILLUSIONS

ظواهر اكتمال الإدراك الحسي استعمل الرؤية والأوهام.

تخيل الآن أن هناك طلباً للوثائق التي تتصل بهذا الموضوع؛ فالموضوع وارد ضمن مصطلحات الدخول إلى لغة التكشيف، ومن ثم فإنه من المعروف أي مجموعة مؤتلفة من المصطلحات يمكن البحث بها. وتحدث أخطاء في التحقيق لأن ناتج الفئة الموسومة «بالرؤية» وتلك الموسومة «بالأوهام»، وهو على وجه التحديد فئة «الأوهام البصرية» أوسع من الفئة المخصصة «ظواهر اكتمال الإدراك الحسي». ورغم ذلك، فإنه لا يمكن أن تحدث أخطاء في الاستدعاء، نظراً لأن مصطلحات الدخول إلى اللغة تكفل للمكتشفين الاطراد في معالجتهم للموضوع، كما تبصر مسئول البحث بالكيفية التي تم بها تكشيف الموضوع.

ولكن، نفترض أنه ليس هناك مصطلح محدد لظواهر اكتمال الإدراك الحسي، وأن الموضوع لم يرد ضمن مصطلحات الدخول إلى اللغة، رغم إدخال وثائق تتناول الموضوع، في النظام، حينئذ يحدث ما يلي:

١ - تجاهل المكشف لبعض الموضوعات؛ فإذا تلقى المكشف وثيقة تتناول ظواهر اكتمال الإدراك الحسي، ولكن ليس كموضوع بؤري، فإنه غالباً ما يتجاهل الموضوع، إذا لم يتوافر مصطلح مخصص سواء في اللغة المقيدة أو في مصطلحات الدخول إلى اللغة.

٢ - حدوث عدم اطراد في التكشيف؛ فبعض المكشفين قد يستعملون «الرؤية والأوهام»، وآخرون يمكن أن يستعملوا تجميعات مؤتلفة متعددة من المصطلحات.

٣ - حدوث أخطاء في الاستدعاء؛ وبعض هذه الأخطاء ناتج عن تجاهل المكشف لبعض الموضوعات، والبعض الآخر راجع لعدم الاطراد (فمسئول البحث لا يعرف كيف صنف الموضوع المطلوب ضمن فئة أكبر، ويعجز عن تغطية كل التجميعات المؤتلفة من المصطلحات، اللازمة لإحراز استدعاء مرتفع).

٤ - حدوث أخطاء إضافية في التحقيق؛ فنظراً لأن مسئول البحث لا يعرف كيف تمت معالجة الموضوع، فإنه يضطر لتجربة عدد كبير من التجميعات المؤتلفة للمصطلحات، وبعضها يمكن أن يؤدي إلى استرجاع عدد كبير من الوثائق غير الصالحة.

ويعني الخطأ في الاستدعاء الناتج عن الافتقار إلى المصطلح المخصص، أن موضوع عملية البحث، أو بعض جوانب هذا الموضوع لم يحظ بالتغطية في مصطلحات الدخول إلى النظام. أما الخطأ في التحقيق الناتج عن الافتقار إلى المصطلح المخصص فيعني أن الموضوع لم يعبر عنه بشكل مخصص بالمصطلحات المقيمة. ولهذا، فإنه لتصحيح أخطاء التحقيق الناتجة عن الافتقار إلى التخصيص، يتعين إدخال مصطلحات أو مجموعات مؤتلفة من المصطلحات تعبر بشكل مخصص عن الموضوع الذي لم يعالج على وجه التخصيص، وذلك في لغة النظام. ولا يتطلب تصحيح أخطاء الاستدعاء استعمال مصطلحات مقيمة مخصصة، وإنما يستلزم التعبير عن الموضوع في مصطلحات الدخول إلى لغة النظام.

وقد حثت أخطاء التحقيق والاستدعاء في جهود الاسترجاع المعتمدة على اللغات المقيمة، مقترنة بالتكلفة في المال والوقت، التي ينطوي عليها إعداد المكانز وصيانتها، حثت البعض على تشجيع الاعتماد على اللغة الطبيعية دون سواها، في مرصد البيانات وفي الاستفسارات (Cleverdon, 1984). وقد حظيت وجهة النظر هذه بالتبني نظراً لما أسفرت عنه تجارب الاسترجاع من نتائج، وأبرزها تلك الدراسات التي أجريت في كرانفيلد بإنجلترا (Cleverdon, Mills and Keen, 1966) والتي تبين منها أن اللغة التي تخضع للحد الأدنى من القيود (التي تم بها فقط تقنين المترادفات ونهايات الكلمات) تعمل بنفس كفاءة أي نوع من اللغات محكمة القيود، وربما كانت أفضل.

وعلى الرغم من ميل الدراسات التي أجريت فيما بعد لتأكيد الصلاحية العالية للغة الطبيعية في الاسترجاع، فقد تبين من نتائج العديد من المشروعات أن كلا من اللغة الطبيعية واللغة المقيمة يكمل بعضهما البعض، أي أن كلا منهما تسترجع وثائق صالحة لا تسترجعها الأخرى. وقد أمكن التوصل إلى هذه النتائج في مجالات موضوعية مختلفة (Calkins, 1982; Katzer et al., 1980) وفي مرصد بيانات تشتمل على كميات متفاوتة من النصوص الحرة، بما في ذلك النصوص الكاملة للوثائق. (Tenopir, 1985; McKinin et al., 1991).

الربط :

مستوى الربط بين المصطلحات من الخصائص الأخرى للغة ، والتي تؤثر في كفاءة الاسترجاع. فإذا نظرنا في بحث عن المعلومات حول الفهارس الكتب ، نجد أن المستفيد إذا أدخل الاستراتيجية « الفهارس والكتب » فإنه يمكن أن يسترجع نوعين مختلفين من الوثائق:

(١) الوثائق التي تتصل بالفهارس التي تعرف بالكتب.

(٢) الوثائق التي تتصل بالفهارس التي تتخذ شكل الكتاب .

إلا أن الفئة الثانية فقط هي المناسبة للطلب. وكانت المشكلات من هذا النوع والناجمة عن العلاقات الغامضة والزائفة بين الكلمات هي السائدة بوجه خاص في نظم المصطلح الواحد Uniterm (الكلمة الواحدة) التي كانت تستخدم في خمسينيات القرن العشرين (Lancaster, 1972) والتي أدت في الواقع إلى تطور المكانز التي تشتمل على بعض العبارات (Holm and Rasmussen, 1961) والتي عززت من دقة اللغة.

وتشتمل لغات الاسترجاع أيضاً على أدوات للدقة أو التحقيق خلاف العبارات، ومن بين هذه الأدوات الربط بين الرؤوس الرئيسة والرؤوس الفرعية أثناء التكشيف (مثل: المكتبات - الاستخدام الآلي - الولايات المتحدة) . ومثل هذه الأداة نجدها في لغات كقائمة رؤوس موضوعات مكتبة الكونجرس *Library of Congress Subject Headings* وقائمة رؤوس الموضوعات الطبية *Medical Subject Headings* . وتشمل أدوات الربط الأخرى الأكثر تعقداً الروابط والأدوار ، وهذه سبق أن تناولناها في الفصل السادس . وعلى ذلك فإنه بالنسبة للموضوع المركب الخاص « بتصميم الطائرات » يمكن أن نجد الأشكال الثلاثة التالية للتعبير :

الطائرات - تصميم (رأس رئيسي / رأس فرعي) AIRCRAFT - DESIGN

تصميم (١) (رابط بين مصطلحين) DESIGN (A)

طائرات (١) AIRCRAFT (A)

تصميم (علاقة دلالية [دور]) DESIGN

طائرات (٤) (بين مصطلحين) AIRCRAFT (4)

وفي حالة الرؤوس الفرعية والروابط يتم الربط بين الجانبين الموضوعيين اللذين يعبر عنهما مصطلحا « الطائرات » و « التصميم » ، في أثناء التكشيف . أما الأدوار فتذهب خطوة أبعد في تحديد العلاقة الدلالية الدقيقة بين المصطلحين ، وهي تبين في هذه الحالة أن الطائرات هي المستهدفة بعملية التصميم .

وكانت الحاجة إلى إدخال الوسائل اللغوية ، وفي مقدمتها الروابط والأدوار ، موضوعاً للتساؤل على ضوء ما انتهت إليه دراسة كرانفيلد من نتائج (Cleverdon, Mills and Keen, 1966) حيث تبين أن مثل هذه الوسائل والأدوات لا لزوم لها . هذا بالإضافة إلى أنه ، كما بين لانكستر (1972) Lancaster ، فإن تكلفة إدخال الأدوات النظامية والدلالية، المتمثلة في عجزها عن استرجاع ما هو مطلوب ، بالإضافة إلى ما يُدفع مقابل استعمالها من أموال، يمكن أن يفوق مزايا تجنب استرجاع بعض الوثائق غير الصالحة بمراحل .

إلا أن الحاجة إلى استعمال العبارات في استراتيجيات البحث، وخاصة في بحث اللغة الطبيعية للنصوص الكاملة، يحظى بالتأييد (Tenopir and Ro, 1990) . فمن الممكن لربط الكلمات المفردة بواسطة العامل « و » أن يؤدي إلى نسبة استدعاء أفضل من تلك التي تنتج عن استعمال العبارات الصريحة، إلا أن ذلك يمكن أيضاً أن يسفر عن مستوى غير مقبول للتحقيق .

البناء :

والخصيصة الثالثة للغة الاسترجاع، والتي تؤثر في كفاءة النظام، هي ما إذا كانت اللغة تنطوي على بناء صريح أم لا . فاللغة المقيدة تفرض بناءً ما ، أما اللغة الطبيعية فلا تفرض أي بناء .

وتساعد اللغة المقيدة مسئول البحث، بتوجيه الانتباه نحو كل المصطلحات التي يمكن أن تدعو إليها الحاجة لإجراء بحث شامل في مجال موضوعي ما . فمن الممكن لأحد المهندسين أن يحتاج إلى إجراء بحث عن الإجهاد في نوع معين من المركبات المعدنية، وعندما يرجع إلى المكنز أو أي نوع آخر من اللغات تحت المصطلح « إجهاد » فإن بناء اللغة ينبغي أن يقوده إلى مصطلحات أكثر تخصيصاً تدل على أنواع مخصصة من

الإجهاد، « كالتشقق » و « الانفجار » مثلاً ، ومن هذه إلى مصطلحات أكثر منها تخصيصاً ، « كالتشقق الهش » و « الانفجار بالضغط » . كما يمكن أن يقوده أيضاً إلى مصطلحات يمكن أن تكون أفضل ، وذلك من تقسيمات هرمية أخرى ، كالمصطلحات التي تعبر عن الظروف التي يمكن أن تؤدي إلى الإجهاد في المركبات المعدنية ، « كالتآكل » ، و « الضغط » ، و « الثبات » ، و « التفتت بالإيدروجين » ، وهكذا . وما لم تفعل ذلك ، فإن اللغة لا تقدم لمسئول البحث أقصى مساعدة ، مما يمكن أن يعرضه لمخاطر إغفال بعض المصطلحات التي قد تدعو الحاجة إليها لإجراء بحث شامل فعلاً .

وتميل اللغة الطبيعية لتعزيز التحقيق ، نظراً لقدرتها على مضاهاة أي مصطلح من مصطلحات الاستفسار ، بما في ذلك المصطلحات باللغة التخصيص التي قد لا ترد في اللغة المقيدة . إلا أن افتقار اللغة الطبيعية للبناء الصريح ، يجعلها مسئولة عن كثير من أخطاء الاستدعاء ، نظراً لأنه يتعين على المستفيد أن يقدم جميع المترادفات ، والاختلافات الهجائية ، وكل ما يمكن أن يتبادر إلى ذهنه من مصطلحات أخرى متصلة بموضوع البحث . وتشمل محاولات تعزيز الاستدعاء في عمليات بحث اللغة الطبيعية ، « اللغة المقيدة تقييداً لاحقاً » (Lancaster, 1986) و « القوائم الاحتياطية » وهي قوائم مخترنة فعلاً بالكلمات والعبارات المتصلة بالموضوعات التي يتم البحث عنها بكثافة ، وذلك في النظم التجارية لاسترجاع المعلومات (Sievert and Boyce (1983) .

جوانب فعالية تكلفة لغة الكشف :

وتحليل فعالية التكلفة قابل للتطبيق على تصميم وتنفيذ لغات الكشف ، إلا أنه من الصعب تطبيقه في هذا المجال ، كما يصعب أيضاً التعبير عن النتائج بشكل ملموس . وكلما ازدادت اللغة إحكاماً ارتفعت تكلفة وضعها وصيانتها . وحجم اللغة من الاعتبارات الاقتصادية المهمة ؛ فكلما زاد عدد المصطلحات التي تشتمل عليها اللغة ، ازداد مدى التخصيص ، وتعززت إمكانات التحقيق الخاصة بالنظام . إلا أن اللغة المقيدة الضخمة عالية التخصيص ، تميل لأن تكون مرتفعة التكلفة في وضعها واستخدامها وتجديدها . ومن ثم ، فإنه يتعين إجراء تحليل دقيق لعينة ممثلة للاستفسارات ، في مرحلة تصميم النظام . وعند النظر في مدى التخصيص في اللغة ، ينبغي مراعاة احتمالات نمو مرصد البيانات ، وأثر هذا النمو على متوسط عدد التسجيلات المسترجعة في عملية

البحث؛ فمن الممكن للمستفيد أن يتحمل نسبة تحقيق ٢٠ بالمئة عندما يكون متوسط ناتج عملية البحث ١٢ تسجيلية ، إلا أنه لا يمكن أن يتحمل هذه النسبة مطلقاً عندما يكون متوسط ناتج عملية البحث ١٢٥ تسجيلية.

وهناك اعتبار وثيق الصلة، وهو الحاجة إلى أدوات تحقيق إضافية ، كالروابط، ومؤشرات الدور، والرءوس الفرعية، ووزن المصطلحات. والهدف من هذه الأدوات هو تعزيز نسبة التحقيق الخاصة بالنظام، بالحد من عدد الوثائق المسترجعة غير المرغوبة في البحث، والنتيجة عن عمليات الربط الزائف ، أو العلاقات غير الصحيحة بين المصطلحات، أو التكشيف بالغ الشمول. واستخدام هذه الأدوات باهظ التكلفة عادة؛ فمن الممكن لمؤشرات الدور ، على وجه الخصوص ، أن تضيف ، وبشكل ملحوظ ، إلى تكاليف التكشيف وصياغة استراتيجيات البحث ، كما يمكن أن تضيف أيضاً إلى تكاليف التجهيز الفعلي لعملية البحث . ونظراً لأنها تزيد من مدى تخصيص اللغة، فإن هذه الأدوات تؤدي دائماً تقريباً إلى الحد من الاطراد في أداء المكشف، كما أنها غالباً ما يكون لها أثر مدمر على الاستدعاء. كذلك تضيف الرءوس الفرعية، التي يمكن أن تقوم في الوقت نفسه مقام الروابط ومؤشرات الدور ، إلى تكلفة التكشيف ، كما تحد من الاطراد في أداء المكشف ، وإن كانت تميل لأن يكون تأثيرها أقل حدة من تأثير مؤشرات الدور. وفي البحث في مراصد البيانات ، تعمل هذه الأدوات على الحد من عدد الوثائق غير الصالحة التي يتعين على المستفيد النظر فيها لكي يعثر على تلك الصالحة، ولا يمكن تبريرها (هذه الأدوات) من وجهة النظر الاقتصادية إلا إذا ثبت أنها أقل تكلفة من الطرق الأخرى، لتحقيق النتائج نفسها للمستفيد النهائي. فمؤشرات الدور، على سبيل المثال، تحد من نوع معين من الاسترجاع الناتج عن العلاقات غير الصحيحة بين المصطلحات، وهو الموقف الذي ترتبط فيه المصطلحات التي تؤدي إلى الاسترجاع ببعضها البعض، ولكن ليس على النحو الذي يريده المستفيد. ويمكن لتحليل فعالية التكلفة أن ينتهي إلى أنه من الأجدي، من الناحية الاقتصادية، عدم استعمال مؤشرات الدور، حيث نقتصد في وقت التكشيف والبحث؛ وقد يترتب على هذا الاقتصاد حدوث بعض الأخطاء في علاقات المصطلحات ببعضها البعض، ومن الممكن التخلص من التسجيلات غير الصالحة التي يتم استرجاعها على هذا النحو، عن طريق عملية تنقية لاحقة للبحث يقوم بها اختصاصيو المعلومات.

ومن المكونات المهمة جداً إلا أنها عادة ما يتم إغفالها في لغة التكشيف، مصطلحات الدخول إلى اللغة ، وهذه عبارة عن تعبيرات باللغة الطبيعية ترد في الوثائق أو الاستفسارات، وتحتل مكانها في اللغة المقيدة الخاصة بالنظام. وعادة ما تتكون لغة الدخول من مصطلحات تعد بالنسبة لأغراض التكشيف والاسترجاع، إما مرادفة لمصطلحات اللغة المقيدة ، أو أكثر تخصيصاً من هذه المصطلحات ، مثل:

HELIARC WELDING use SHIELDED ARC WELDING

لحام الأقواس الشمسية استعمل لحام الأقواس المدرعة

وعلى الرغم من أن لغة الدخول المطولة يمكن أن تكون باهظة التكلفة في إعدادها وتجديدها، فإنها يمكن أن ترفع من مستوى الأداء بشكل ملحوظ ، وذلك بالحد من أخطاء الاستدعاء ، وخاصة في نظم الاسترجاع الضخمة . كذلك يمكن لمثل هذه اللغة أن تجعل النظام أكثر فعالية من وجهة نظر التكلفة، بمرور الوقت، وذلك بالحد من العبء الفكري الواقع على عاتق كل من المكشفين ومسئولي البحث . ف لغة الدخول في الواقع مجموعة من تسجيلات القرارات الفكرية التي اتخذت أساساً من جانب المكشفين. وما لم يتم تسجيل القرار الفكري الذي يتخذه المكشف ، بتكشيف الموضوع س تحت المصطلح ص، فسوف يكون من الضروري اتخاذ القرار ثانياً، ولكن ليس بنفس طريقة التعبير عن النتائج بالضرورة (ومن هنا يأتي عدم الاطراد) من جانب مكشفين آخرين، أو من جانب المكشف نفسه في وقت لاحق. هذا بالإضافة إلى أنه سوف يتعين على مسئول البحث في النظام، أن يتخذوا أيضاً قرارات فكرية، قد لا تتفق مع تلك التي اتخذها المكشفون، عندما يكونون بصدد البحث عن الإنتاج الفكري المتصل بالموضوع س. وكلما ازداد حجم لغة الدخول، انخفض عدد القرارات الفكرية التي يتطلب الأمر اتخاذها من جانب المكشفين ومسئولي البحث، وبذلك يتحقق الاقتصاد في وقت التكشيف والبحث. وكلما ارتفع مستوى الاطراد في التكشيف، تحسن استدعاء النظام، وربما أيضاً انخفض المستوى المهني للعاملين، والذي تدعو الحاجة إليه في عملية التكشيف.

التكشيف :

كما سبق أن بينا في الفصل الخامس، فإنه يمكن لتكشيف الوثائق أن يتم بطرق

مختلفة، بما في ذلك التكشيف بواسطة البشر، سواء بالتعيين أو بالاقتطاف، أو ببساطة بتبني كلمات اللغة الطبيعية وعباراتها الواردة في الحقول النصية (التي تشمل العنوان ، والمستخلص ، والنص الكامل للوثيقة، والمصطلحات الكشفية) . وينطوي التكشيف البشري على عملية انتقاء ؛ حيث يتم اقتطاف كلمات أو عبارات معينة للتعبير عن الموضوع الذي تتناوله الوثيقة ، أو التعبير عن هذا الموضوع بمجموعة من المصطلحات المنتقاة من لغة مقيدة. ويمكن في بعض الإجراءات الآلية أن تنطوي العملية على انتقاء جميع الكلمات التي لم ترد في « قائمة استبعاد » تم إعدادها مسبقاً.

وعلى ذلك فإنه يتم اتخاذ عدة قرارات حول تكشيف الوثائق ، تؤثر في كفاءة الاسترجاع. وتشمل هذه القرارات :

١ - قرارات السياسة :

أ - مدى الشمول في التكشيف ، أي مدى الإحاطة أو الإلمام بكل الموضوعات البارزة التي تتناولها الوثيقة، والتعبير عنها من جانب المكشف.

ب - مدى التخصيص في استعمال المصطلحات ، أي التعبير عن الموضوعات التي أمكن التحقق منها بلغة تضاهي هذه الموضوعات في التخصيص .

٢ - قرارات المكشف (الدقة) :

أ - التعبير عن جميع الموضوعات التي ينبغي التعبير عنها دون سواها.

ب - اختيار المصطلحات الملائمة.

الشمول :

يرتبط أهم قرارات السياسة المتصلة بالتكشيف بالشمول . لنفترض أن هناك وثيقة تتناول ستة موضوعات (أ ، ب ، ج ، د ، هـ ، و) فإذا كانت جميع الموضوعات الستة قد أمكن التحقق منها في مرحلة التحليل الموضوعي أثناء التكشيف ، وتم التعبير عنها بالمجموعات المؤلفة المناسبة من المصطلحات الكشفية ، حينئذ يكون المكشف شاملاً تماماً في تكشيف هذه الوثيقة. فإذا كانت جميع الموضوعات الستة قد كشفت، فإنه يمكن استرجاع الوثيقة بصرف النظر عن طلب أي من هذه الموضوعات أو

المجموعة المؤتلفة من الموضوعات . ومن ثم ، فإن مستوى الشمول المرتفع في الكشف يميل لضمان الاستدعاء المرتفع . وبقدر ما ينخفض مستوى الشمول تتناقص إمكانات الاستدعاء . فإذا كان الموضوع « و » لم يتم التحقق منه عند كشف الوثيقة ، فإن هذه الوثيقة لا يمكن أن تسترجع استجابة لطلب للإنتاج الفكري حول الموضوع « و » ما لم تكن المصطلحات المستعملة للتعبير عن هذا الموضوع ، قد حدث أن كانت متصلة هرمياً ، أو على أي نحو آخر بالمصطلحات المستعملة للتعبير عن « أ و ب و ج و د و هـ » . وبقدر ما ينخفض مستوى الشمول تدريجياً نتيجة لتتابع تجاهل المزيد من الموضوعات ، تتناقص إمكانات الاستدعاء تبعاً لذلك . وتنطبق هذه الظاهرة على جميع الوثائق المكشوفة على قدم المساواة . فإذا كانت جميع الوثائق قد كشفت بأقصى درجات الشمول ، فإنه من الممكن بلوغ أقصى إمكانات الاستدعاء .

وعلى الرغم من أن ارتفاع مستوى الشمول يميل لضمان ارتفاع الاستدعاء فإنه يميل أيضاً لخفض التحقيق . فإذا أمكن بالنسبة لكل وثيقة تدخل في النظام التعرف على كل ما بها من موضوعات قابلة للكشف ، أو على قطاع كبير من هذه الموضوعات على الأقل ، فإن كثيراً من الموضوعات التي يتم كشفها ، قد لا تحظى إلا بقدر ضئيل من الاهتمام في محتوى الوثائق ، أو يمكن أن تعالج بشكل سطحي عارض . وتبعاً لذلك فإنه يمكن استرجاع الوثائق استجابة لطلبات ، لا تشتمل بالنسبة لها إلا على قدر ضئيل من المعلومات .

وربما كانت الوثيقة ذات الموضوعات الستة (أ ، ب ، ج ، د ، هـ ، و) تتناول بعض ظواهر الديناميكا الهوائية ، وربما كان هناك أيضاً ذكر لأسلوب رياضي معين ، وليكن أسلوب بولهاوزن Pollhausen باعتباره قابلاً للتطبيق في الحسابات الخاصة بإحدى هذه الظواهر . ونفترض أيضاً أن هذا الأسلوب الرياضي قد أمكن التحقق منه في مرحلة التحليل الموضوعي في الكشف ، وترجم إلى المصطلحات الكشفية الملائمة ، وكان هذا الأسلوب هو الموضوع « و » في الواقع . وبالنسبة للطلب الذي يحتمل أن يكون نادراً ، والذي يرغب فيه أحد الباحثين استرجاع جميع الوثائق التي تتناول تطبيق أسلوب بولهاوزن ، فإن هذه الوثيقة تعد صالحة وينبغي أن تسترجع . ويؤكد ارتفاع مستوى الشمول في الكشف جدواه فعلاً بالنسبة لهذا الطلب مرتفع الاستدعاء .

ولكننا إذا نظرنا في طلب آخر ، ربما كان أكثر شيوعاً ، حيث يرغب المستفيد في الوثائق التي تتناول بالوصف أسلوب بولهاوزن وكيفية تطبيقه. فهذا المستفيد لا يريد سوى الوثائق الأساسية حول الأسلوب ؛ فالوثائق التي تكتفي بذكر الأسلوب عرضاً ليست صالحة لهذا الطلب. إلا أنه يمكن للبحث في مجموعة الديناميكا الهوائية ، المكشوفة على مستوى عالٍ من الشمول ، أن يميل لاسترجاع عدد كبير من الوثائق التي لا تقدم أكثر من ذكر الأسلوب عرضاً. ويؤدي ارتفاع مستوى الشمول في هذه الحالة إلى استرجاع عدد كبير من الوثائق غير المرغوبة.

وهكذا ، تميل أخطاء الاستدعاء للوقوع عندما يكون الشمول منخفضاً. أما أخطاء التحقيق فتقع عندما يرتفع مستوى الشمول ، بالنسبة لعمليات البحث التي لا يرغب المستفيد منها إلا في الوثائق التي تتناول الموضوع بشكل أساس. أضف إلى ذلك أن ارتفاع مستوى الشمول يميل للحد من التحقيق ، نظراً لأنه كلما زاد عدد الموضوعات التي يتم التحقق منها في الكشف ، وزاد عدد المصطلحات الكشفية المستعملة للتعبير عن هذه الموضوعات ، ازدادت احتمالات الارتباطات الزائفة في عملية البحث. ومن ثم ، فإنه في الوثيقة ذات الموضوعات الستة (أ ، ب ، ج ، د ، هـ ، و) والتي ربما يرتبط فيها «أ» ب «ب» ، و «ج» ب «د» ، و «هـ» ب «و» ، هناك احتمال استرجاع الوثيقة خطأ استجابة لأي من اثني عشر طلباً ثنائي الموضوع : «أ» مرتبطاً ب «ج» ، و «أ» مرتبطاً ب «د» ، و «ب» مرتبطاً ب «ج» ، و «ب» مرتبطاً ب «د» ، و «ج» مرتبطاً ب «د» ، و «ج» مرتبطاً ب «هـ» ، و «د» مرتبطاً ب «هـ» ، و «د» مرتبطاً ب «و» ، و «هـ» مرتبطاً ب «و» ، و «و» مرتبطاً ب «ج» ، و «و» مرتبطاً ب «د» ، و «و» مرتبطاً ب «هـ» ، و «و» مرتبطاً ب «و» . أما إذا كان كل من الموضوعات الستة قد تم التعبير عنه بعدد من المصطلحات الكشفية المتميزة غير المرتبطة ببعضها البعض ، فإن احتمالات الارتباطات الزائفة ، على مستوى المصطلح الكشفية ، تتضاعف بشكل ملحوظ.

ولننظر الآن في موقف أعدت فيه الكشافات على أدنى مستوى للشمول ، حيث تم كشف كل وثيقة تحت موضوع واحد فقط ، وهو الموضوع المحوري الذي يحظى بالمعالجة في كل حالة. فاحتمالات الاستدعاء الخاصة بالكشاف في غاية التدني ، حيث لا يمكن للوثيقة أن تسترجع استجابة لطلب موضوع دون الموضوع البؤري للمناقشة. ولكن في كل مرة يتم فيها استرجاع وثيقة ما عند البحث في مرصد البيانات ، فإن هذه الوثيقة تميل لأن تكون الوثيقة المرغوبة ، لأنها تتناول أساساً موضوع الطلب. هذا

بالإضافة إلى أنه بالتكشيف بأدنى مستويات الشمول يتم التخلص نهائياً من احتمالات الارتباطات الزائفة على مستوى الموضوع، كما يتم أيضاً الحد وبشكل ملحوظ من الارتباطات الزائفة على مستوى المصطلح.

ومستوى الشمول المتبع في التكشيف أحد قرارات السياسة التي يقرها مديرو نظام الاسترجاع. وهو لا يتوقف على خصائص لغة التكشيف، نظراً لأنه يتعين التسليم بأن لغة التكشيف المستخدمة ملائمة للمجالات الموضوعية التي تحظى بالاهتمام في وثائق المدخلات، وأنه من الممكن ترجمة أي موضوع تتناوله هذه الوثائق إلى لغة النظام، حتى وإن تم ذلك على مستوى أكثر تعميمًا. هذا بالإضافة إلى أن مستوى الشمول لا يخضع لسيطرة المكشف، ذلك لأنه عند اتخاذ هذا القرار الخاص بالسياسة، يتعين على مديري النظام محاولة تحديد مستوى أمثل للشمول، المستوى الذي يمكن أن يلبي الغالبية العظمى من الطلبات التي يتلقاها مرفق المعلومات. ومن الممكن لمستوى الشمول أن يكون متصلاً بنوعية الوثائق؛ حيث يمكن تكشيف التقارير التقنية الداخلية بمستوى شمول أعلى من غيرها من أنواع الوثائق، وتكشيف بعض الدوريات بمستوى شمول أعلى من دوريات أخرى، وهكذا. وفي مقابل ذلك يمكن للسياسة بالنسبة للشمول أن تتفاوت تبعاً لاختلاف الموضوعات المخصصة التي تحظى بالاهتمام في مجموعة الوثائق. وشمول التكشيف بوجه عام، هو أهم العوامل المؤثرة في الاستدعاء الذي يمكن بلوغه في نظام معين، كما أن التخصيص في اللغة هو أهم العوامل المؤثرة فيما يمكن إنجازه من تحقيق في هذا النظام (Lancaster and Mills (1964). وقد تناول عدد من الباحثين، أحدثهم بويس وماكلين (Boyce and McLain (1989 وشو Shaw (1990 a, b) وبيرجن (Burgin (1991، بالدراسة أثر الشمول على كفاءة الاسترجاع.

وهناك اختلافات كبيرة في الشمول بين مراصد البيانات التي يمكن التعامل معها على الخط المباشر. ويقارن راينغ وبيفر (Reich and Biever (1991 بين أجريكولا AGRICOLA ومستخلصات كاب CAB ABSTRACTS في هذا الصدد. وما يثير الاهتمام أنه على الرغم من أن متوسط عدد المصطلحات المخصصة للوثيقة كان مماثلاً (بناءً على عينة من الوثائق) فقد كانت هناك أوجه اختلاف كثيرة بين مرصدي البيانات عند النظر في بعض الوثائق كل على حدة.

ويميل الكشف الآلي لحقول اللغة الطبيعية لأن يسفر عن تسجيلات للوثائق أكثر شمولاً . ومن ثم ، فإنه لا عجب أن يتبين من بعض الدراسات أن البحث في مرصد البيانات المكشفة على هذا النحو يعزز الاستدعاء (Markey, Atherton and Newton (1980) . ووفقاً للمبدأ نفسه، تعمل اللغة الطبيعية على دعم التحقيق (Carrow and Nugent, (1977). ويرى سفينونوي (1986) Svenonius أن هذا موقف معقد ، لم تتضح أبعاده بعد :

« مما يدل على الحاجة إلى البحث انتهاء البحوث إلى نتائج متضاربة ؛ فالنتائج التي انتهت إليها دراسة ماركي وأثرتون ونيوتن (1980) Markey, Atherton and Newton من أن الاسترجاع اعتماداً على اللغة الطبيعية يسفر عن استدعاء أعلى وتحقيق أدنى مما يسفر عنه الاسترجاع اعتماداً على اللغة المقيدة ، لا يتعارض مع ما انتهت إليه الدراسات السابقة من نتائج فحسب، وإنما يناقض «الحكمة التقليدية» أيضاً؛ فالحكمة التقليدية تقضي بأن مصطلحات اللغة الطبيعية تسهم في تعزيز التحقيق نظراً لأنها أكثر تخصيصاً وأكثر حداثة من مصطلحات اللغة المقيدة ».

التخصيص :

من سياسات الكشف الأخرى التي تؤثر في الاستدعاء والتحقيق، ما إذا كان من الممكن للمصطلحات المقابلة لكل موضوع أن تكون متساوية في التخصيص ، أو ما إذا كان من الضروري أيضاً تعيين مصطلح إضافي على مستوى تخصيص أعلى . لنفترض على سبيل المثال، أنه في مرصد البيانات « س » استعمل أحد المكشفين الواصفة «مكتبات عامة»، وهي مساوية في التخصيص للموضوع الذي أمكن التحقق منه، هذا بالإضافة إلى أنه بناء على سياسة التخصيص في تعيين المصطلحات ، كان هذا هو المصطلح الوحيد الذي يمكن تعيينه للتعبير عن هذا الموضوع. إلا أنه في مرصد البيانات ص تم أيضاً تعيين واصفة إضافية على المستوى الأعلى التالي، للتعبير عن هذا الموضوع، وربما عين المكشف في هذه الحالة المصطلح الإضافي « المكتبات » . فإذا افترضنا أن هناك في كل واحد من هذين المرصدين ٥٠ وثيقة تتناول المكتبات العامة على وجه التخصيص، و ٥٠ وثيقة عن المكتبات بوجه عام، ومن الممكن في مرصد البيانات «س» كشف خمسين وثيقة تحت المصطلح « المكتبات العامة » وخمسين وثيقة تحت المصطلح « المكتبات »، أما في مرصد البيانات « ص » فإنه فيما يتعلق بالوثائق المئة نفسها، ربما تم كشف خمسين وثيقة تحت المصطلح « المكتبات العامة » وكشف الوثائق المئة جميعاً تحت المصطلح « المكتبات ». ويصور الشكل رقم (٥٦) هذا الموقف.

ويؤدي تعيين مصطلح واحد فقط مخصص في مستوى تخصيص الموضوع المكشف نفسه، إلى دعم التحقيق، نظراً لأنه في أي عملية بحث لن تسترجع سوى الوثائق الواقعة في هذا المستوى من التخصيص. وربما يكون هناك، بالطبع، بعض الفاقد في الاستدعاء؛ حيث يمكن للوثائق المكشوفة على مستوى أكثر تعميماً أن تشتمل على معلومات يمكن للمستفيد أن يراها صالحة. هذا بالإضافة إلى أنه نظراً لأن التكشيف لا يمكن أن يكون كاملاً، فإنه يمكن تكشيف بعض الوثائق تحت المصطلح الأعم، بينما كان ينبغي تكشيفها تحت المصطلح الأكثر تخصيصاً.

| مرصد البيانات | المصطلحات المستعملة في تكشيف ٥٠ وثيقة تتناول المكتبات العامة على وجه التخصيص | المصطلحات المستعملة في تكشيف ٥٠ وثيقة تتناول المكتبات بوجه عام | عدد الوثائق المسترجعة في البحث عن المكتبات بوجه عام | عدد الوثائق المسترجعة في البحث عن المكتبات العامة |
|---------------|--|--|---|---|
| س | المكتبات العامة | المكتبات | ٥٠ | ٥٠ |
| ص | المكتبات العامة المكتبات | المكتبات | ١٠٠ | ٥٠ |

الشكل رقم (٥٦) : تعيين افتراضي للمصطلحات الكشفية في مرصد البيانات الذي يكشف وفقاً للمصطلح المخصص دون غيره (س) ومرصد (ص) لا يلتزم بهذه القاعدة

وحيثما يعين المكشفون المصطلحات على كل من المستوى المخصص والمستوى العام، يميل الاستدعاء للارتفاع، نظراً لأن البحث على المستوى العام (كل ما يتصل بالمكتبات على سبيل المثال) سوف يسترجع تلقائياً وثائق على مستويات أكثر تخصيصاً (حول المكتبات العامة على سبيل المثال). ووجه القصور هنا بالطبع، أنه يصبح من الصعب بمكان التمييز بين الوثائق التي تتناول الموضوع بوجه عام وتلك التي تتناول الأوجه الأكثر تخصيصاً.

وتلتزم معظم مرافق المعلومات « بقاعدة المدخل المخصص دون سواه » (Wilson, 1979) والاستثناء البارز في هذا الصدد مكتبة الكونجرس، التي تعين لبعض فئات الوثائق رأسين، أحدهما مخصص والآخر عام. ومن ثم، فإن ترجمة أحد المؤلفين الأمريكيين المعاصرين، يمكن أن يعين لها لا اسم الشخص فحسب، وإنما فئة البشر التي ينتمي إليها الفرد أيضاً، (المؤلفون الأمريكيون - القرن العشرون - تراجم، على سبيل المثال).

على الرغم من أن السياسات الخاصة بالشمول والتخصيص في تعيين المصطلحات الكشفية لا تخضع لسيطرة المكشف، فإن نوعية التكشيف ودقته تدخل في صميم نطاق سيطرته. وهناك نوعان من أخطاء المكشفين :

(١) إغفال مصطلحات لا غنى عنها لوصف موضوع مهم تتناوله الوثيقة.

(٢) استعمال مصطلحات تبدو غير ملائمة لموضوع الوثيقة.

وتؤدي حالات الإغفال عادة إلى أخطاء الاستدعاء ، أما المصطلحات غير الصحيحة أو غير المخصصة فيمكن أن تكون سبباً في أخطاء التحقيق (حين يستعمل مسئول البحث المصطلح في استراتيجية ما ، ويسترجع وثيقة غير صالحة) أو في أخطاء الاستدعاء (حين يستعمل مسئول البحث المصطلحات الصحيحة ، وتضيع وثيقة مرغوبة، لأنها وضعت تحت وسيمة غير صحيحة) .

وينبغي التمييز بين أخطاء الاستدعاء الراجعة إلى إغفال المكشف للمصطلحات ، وتلك التي ترجع إلى افتقار التكشيف إلى الشمول. وتقع أخطاء الإغفال حيثما يتجاهل التكشيف موضوعاً يبدو محورياً في محتوى الوثيقة، على الرغم من أنه من المهم بمكان تغطيته حتى وإن لم يكن التكشيف شاملاً. وحينما يكون مستوى الشمول في التكشيف منخفضاً بناء على سياسة النظام، فإن الجوانب الموضوعية التي تتناولها الوثائق بشكل سطحي أو على نحو هامشي، لا تحظى بالتغطية في التكشيف. فمثل هذه الموضوعات ليست محورية ولا أساسية بالنسبة للوثائق، ومن ثم فإنها تستبعد لصالح موضوعات أخرى، نتيجة للسياسة العامة المتصلة بمتوسط عدد المصطلحات التي يتم تعيينها للوثيقة.

ويمكن ، للأسف ، إذا ما أغفل مصطلح مهم في التكشيف، أن تظل الوثيقة دون استرجاع، في عمليات البحث التي كان من الممكن أن تكون بالنسبة لها في غاية الصلاحية. أضف إلى ذلك، أن هذا النوع من الأخطاء، على الرغم من احتمال الكشف عنه في برنامج التقييم، يمكن أن يظل مستتراً في التشغيل العادي للنظام. ومن الممكن بالطبع توقع عدد معين من حالات الإغفال في التكشيف، نتيجة لضغوط جداول الإنتاج الصارمة. ويمكن لحالات الإغفال هذه أن تؤدي إلى حدوث أخطاء في النظام أكثر من تلك

الناجمة عن استعمال المصطلحات غير الملائمة. وإذا تعرض عمل المكشف للفحص أو التدقيق (المراجعة) من قبل مكشف آخر، فإن المصطلحات غير الملائمة تميل للظهور بوضوح، بحيث يكون من السهل تصحيحها في عملية المراجعة. أما حالات الإغفال فلا يمكن اكتشافها بيسر، لأن ذلك قد يتطلب مزيداً من التدقيق في فحص الوثيقة من جانب المراجع.

جوانب فعالية التكلفة في الكشف:

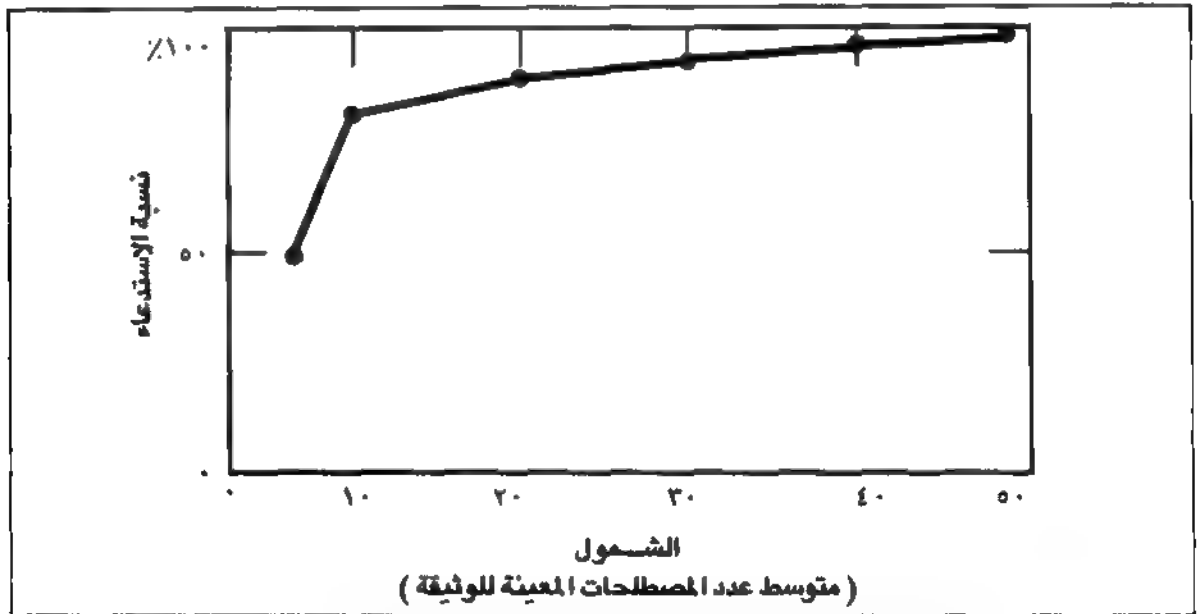
من الممكن تبرير الاستغناء عن الكشف بواسطة البشر بما يتحقق من اقتصاد في تكلفة المدخلات، فضلاً عن الفورية في بث المعلومات إلى المستفيدين. إلا أنه من الممكن أيضاً للاحتفاظ بعنصر الكشف بالتعيين أن يكون له ما يبرره نظراً لما يضيفه على المجموعة من تنظيم، ييسر صياغة استراتيجيات البحث الفعالة ويقتصد في الوقت المستنفد في صياغة هذه الاستراتيجيات. كما يمكن تبريره أيضاً لأنه تبين أن اللغة الطبيعية واللغة المقيدة تكمل كل منهما الأخرى كأدوات للاسترجاع.

وإذا افترضنا تبرير عنصر الكشف بالتعيين، فإن هناك عدداً من الاعتبارات الاقتصادية الجديرة بالدراسة، فيما يتعلق بسياسات الكشف وإجراءاته:

- ١ - مقدار الوقت المستنفد، في المتوسط، في كشف الوثيقة.
- ٢ - مستوى الشمول المتبع في الكشف، أي عدد المصطلحات الكشفية التي يتم تعيينها، في المتوسط، للوثيقة.
- ٣ - المستوى المهني للعاملين في الكشف.
- ٤ - الحاجة إلى أسلوب لمراجعة الكشف.

وربما كان من الممكن لسياسة الكشف بالغة الصعوبة في إقرارها، أن تنطوي على أنسب مستوى للشمول يمكن اتباعه. فكلما ازداد الشمول في الكشف ازدادت احتمالات ارتفاع الاستدعاء في النظام، ولكن مع انخفاض التحقيق. ومن الممكن للمستوى الأمثل للشمول في الكشف أن يتحقق في بيئة متفردة من الوثائق ولغة الكشف والطلبات. ويكفل تحليل فعالية التكلفة لنا القدرة على اكتشاف هذا المستوى الأمثل، أي الحد الذي تصبح بعده إضافة المزيد من المصطلحات غير مجدية على الإطلاق. فإذا افترضنا أن هناك مجموعة معينة من الوثائق كشفت بمعدل ١٥ مصطلحاً

للوثيقة، وتم اختبار كفاءة الاسترجاع في النظام اعتماداً على عينة ممثلة من الطلبات قوامها ٥٠ طلباً. وربما يتبين، باتباع استراتيجيات البحث المثل، أن النظام بالنسبة لهذه المجموعة بعينها من الطلبات، وعلى مستوى الكشف المتبع، يعمل بنسبة استدعاء ٧٢ بالمئة. وعن طريق بعض تجارب الكشف الإضافية، والتي يتم فيها المحافظة على ثبات كل من الطلبات، واستراتيجيات البحث، ولغة الكشف، يتبين أن الزيادة في الشمول إلى معدل ٢٠ مصطلحاً للوثيقة، تؤدي إلى ارتفاع نسبة الاستدعاء إلى ٩٠ بالمئة، إلا أن متوسط مستوى الشمول ينبغي أن يرتفع إلى ٣٥ مصطلحاً للوثيقة لبلوغ نسبة استدعاء ٩٥ بالمئة. وفي ظل هذه الظروف، ربما كان من الممكن إقرار عشرين مصطلحاً في المتوسط للوثيقة، كمستوى أمثل يمكن الالتزام به في هذه البيئة من الوثائق والطلبات ولغة الكشف، بالذات. فهناك بعد مستوى العشرين مصطلحاً ما يدل على تناقص العائدات، نظراً لأنه لتحقيق تحسن بنسبة ٥ بالمئة في إجمالي الاستدعاء، يتعين الارتفاع بمستوى الشمول إلى ٣٥ مصطلحاً، مما يمكن أن يؤدي إلى زيادة تكاليف الكشف بشكل ملحوظ. واعتماداً على النتائج المستخلصة من هذه الدراسة يمكن توقع منحني الأداء من مستوى الشمول في مقابل نسبة الاستدعاء (الشكل رقم ٥٧).



الشكل رقم (٥٧) : منحني أداء افتراضي لمستوى الشمول مقابل نسبة الاستدعاء.

واعتماداً على رسم بياني من هذا النوع يمكن تحديد نقطة حدية للشمول في الكشف، وهي النقطة التي لا يمكن بعدها لإضافة المزيد من المصطلحات الكشفية، أن تؤدي إلى تغير يذكر في احتمالات استدعاء النظام، رغم ما تؤدي إليه من زيادة كبيرة في

تكاليف المدخلات. فإذا جمعت عينة من الاستفسارات الموضوعية التي تمثل الاحتياجات الفعلية للمستفيدين المحتملين ، وإذا أمكن التعرف على وثائق معينة باعتبارها صالحة لهذه الاستفسارات ، فإن هذا النوع من تحليل فعالية التكلفة يمكن ، بل ينبغي أن يتم في مرحلة تصميم النظام.

ومن الممكن التعامل مع المشكلة نفسها بأسلوب آخر ، وذلك بدراسة أثر تفاوت الوقت المستنفد في الكشف على عملية الكشف: فكلما ازداد الوقت المستنفد في الكشف ازداد عدد المصطلحات التي يتم تعيينها في المتوسط، إلا أنه يمكن أن يكون هناك هبوط مؤكد في عدد المصطلحات التي يتم تعيينها بعد استنفاد قدر قليل نسبياً من الوقت، حيث يمكن على سبيل المثال تعيين عشرين مصطلحاً في الدقائق العشر الأولى، وخمسة مصطلحات إضافية في الدقائق العشر التالية، ومصطلحين آخرين في الدقائق العشر التي تليها. فإذا كان تعيين ٩٠ بالمئة من المصطلحات المناسبة يتم في الدقائق العشر الأولى من وقت الكشف، بينما يتطلب العثور على العشرة بالمئة الباقية من المصطلحات التي يمكن الحكم بأنها مناسبة، ثلاثين دقيقة إضافية، فإنه من الممكن أن ننتهي باطمئنان إلى أن عشر دقائق للوثيقة هو المتوسط الأمثل لوقت الكشف، وتحديد حصص الكشف على هذا الأساس.

ومن الممكن إجراء تحليل لفعالية تكلفة الوقت المستنفد في الكشف، بإجراء تجربة مماثلة لتلك التي تعرضنا لها بالنسبة لتحديد المستويات المثلى للشمول. ونظراً لأنه عادة ما يتم بالفعل تعيين المزيد من المصطلحات تبعاً لزيادة وقت الكشف، فإننا نقيس الشيء نفسه في الأساس.

والتنبؤ بأثر التفاوت في مستويات الشمول أو في الوقت المستنفد في الكشف على التحقيق في نظام استرجاع المعلومات، أمر أكثر صعوبة، على الرغم من أنه يعد عنصراً أساسياً لا غنى عنه في التحليل الحقيقي لفعالية التكلفة. ويتطلب هذا التنبؤ إجراء تجربة ضخمة ، تشمل بالإضافة إلى الكشف مجموعة من الوثائق، يستقر الرأي في البداية على أن كلاً منها تعد صالحة بالنسبة لواحد أو أكثر من الطلبات المسجلة، ويقوم مسئول التقييم بتكشيف عينة عشوائية من الوثائق الإضافية في المجال الموضوعي العام نفسه. وعند مضاهاة استراتيجيات البحث الخاصة بطلبات الاختبار مقابل المصطلحات الكشفية المعينة لكل من الوثائق الصالحة والوثائق غير الصالحة، يمكن قياس كل من

الاستدعاء والتحقيق لمجموعة الاختبار، ثم يتم تقدير التحقيق استقرائياً، لنظام مرصد البيانات بأكمله. وقد استخدم هذا الأسلوب بنجاح في تقييم نظم الاسترجاع التجريبية.

ويمكن للتفاوت في زمن الكشف أن يؤثر لا في متوسط الشمول في الكشف فحسب، وإنما في دقة الكشف أيضاً. ولهذا، فإن تحليل فعالية التكلفة المتصل بالوقت المستنفد في الكشف ينبغي أن يأخذ في الحسبان أثر الوقت المسموح به على الدقة وعلى الشمول أيضاً. وكلما ازدادت ضغوط الإنتاج ازدادت أخطاء الكشف التي يحتمل أن تقع. ويمكن لأخطاء الكشف أن يكون لها أثر مضاعف على الاستدعاء في نظم الاسترجاع المعتمدة على الربط اللاحق. ومن الممكن توضيح ذلك بمثال بسيط؛ نفترض أنه في كشف مجموعة معينة من الوثائق تم تعيين المصطلح «أ» لتسعين بالمئة من الوثائق التي كان ينبغي أن يعين لها بإجماع الآراء، وتعيين المصطلح «ب» لخمس وثمانين بالمئة من الوثائق التي كان ينبغي أن يعين لها، وتعيين المصطلح «ج» لخمس وسبعين بالمئة من الوثائق الملائمة، فإن البحث بالمصطلح «أ» وحده يسترجع ٩٠ بالمئة من الوثائق التي كان ينبغي استرجاعها، أي أن نسبة الاستدعاء في وثائق «أ» ٩٠ بالمئة. إلا أنه بالبحث بكل من «أ و ب» يمكن للاستدعاء أن ينخفض إلى ٧٦,٥ بالمئة (٩٠ بالمئة \times ٨٥ بالمئة) بينما يمكن للبحث بكل من «أ و ب و ج» أن يسفر عن نسبة استدعاء قدرها ٥٧,٤ بالمئة فقط (٩٠ بالمئة \times ٨٥ بالمئة \times ٧٥ بالمئة) من الوثائق التي تنتمي، كما ينبغي للمجموعة «أ، ب، ج». ومن ثم، فإنه للتعرف على أثر الوقت المستنفد في الكشف على دقة الكشف، من وجهة نظر فعالية التكلفة، ينبغي على مسئول التقييم مراعاة هذا التأثير المضاعف على كفاءة الاسترجاع. وعلى مسئول التقييم أن يكون قادراً على تقدير معدل الدقة في تعيين المصطلحات في لغة النظام بأكملها، وربط ذلك بمتوسط مستوى الربط المتبع في عمليات البحث. وقد حظي تأثير مختلف أنواع أخطاء الكشف على فعالية البحث المعتمد على الربط اللاحق باستعمال عدة استراتيجيات بحث، بالاهتمام تفصيلاً، من جانب كل من براينت وكنج وتراجنو (Bryant, King and Terragno (1963). ويصف كنج (King (1965, 1967) أحد تطبيقات هذه الأساليب، كما طور نموذجاً مفيداً لتحويل البيانات الخاصة بالاطراد في أداء المكشف إلى بيانات خاصة بدقة المكشف، والعكس.

والقضية المرتبطة بذلك ارتباطاً وثيقاً هي ما إذا كانت الحاجة تدعو إلى عملية مراجعة للكشف أم لا، أي ما إذا كان من الضروري مراجعة ما يقوم به أحد المكشفين

من جانب شخص آخر أكثر تمرساً عادة أم لا . وما إذا كانت عملية المراجعة لها ما يبررها بناء على أسس فعالية التكلفة أم لا ، أمر يتوقف على :

- ١ - مقدار الأخطاء الواردة في الكشف الذي لم يراجع .
- ٢ - مقدار الأخطاء التي يتم تصويبها بعملية المراجعة.
- ٣ - الأثر المقدر لأخطاء الكشف المراجع وغير المراجع، على كفاءة الاسترجاع.
- ٤ - تكلفة المراجعة.

ويمكن لبعض التجارب البسيطة غير المكلفة ، والتي يتم فيها إجراء عملية المراجعة والتعرف على ما تستنفده من وقت ، وذلك على مجموعة من الوثائق المكشوفة والتي يشتمل بعضها على أخطاء معروفة ، يمكن لهذه التجارب أن تسفر عن تحديد معدل التصويب المتوقع لعملية المراجعة فضلاً عن تكاليفها. وينبغي أن يختلف المراجعون حتى يكون من الممكن التعرف على معدل أداء المراجع، فضلاً عن دراسة كفاءة مختلف فئات العاملين . ومن الممكن بعد ذلك ربط النسبة المئوية للحد من الأخطاء بالتأثير المقدر للمراجعة على كفاءة الاسترجاع. ويمكن لمرقق المعلومات، نتيجة لهذا التحليل أن يكون قادراً على القول بأنه من الممكن لعملية المراجعة التي تتكلف س من الدولارات في العام أن تصوب ص بالمئة من الأخطاء التي تقع في الكشف ، وأنه من الممكن لذلك أن يكون له أثر محدد على نسب الاستدعاء والتحقيق في النظام.

وهناك جانب آخر لعملية الكشف يمكن إخضاعه لتحليل فعالية التكلفة، وهو مستوى القوى البشرية اللازمة للقيام بالكشف . فمن الممكن الحد من تكاليف الكشف بشكل ملحوظ إذا أمكن الاعتماد بنجاح على مكشفين ليسوا على مستوى عالٍ من التأهيل ويتقاضون مرتبات منخفضة. وقد تحولت مؤسسة ضخمة واحدة على الأقل من النظام الذي كان الكشف يتم فيه اعتماداً على محللين متمرسين على مستوى عالٍ، جميعهم من الحاصلين على مؤهلات جامعية، إلى نظام يتم فيه معظم الكشف اعتماداً على عاملين من غير الحاصلين على مؤهلات جامعية. وهناك افتراض آخر أحياناً ما يساق، وهو أن المكشف الجيد بحاجة لأن يكون مؤهلاً في المجال الموضوعي للوثائق التي يتعامل معها. ولم يحدث لهذا الافتراض أن تأكد بشكل قاطع ، بل إن معظم ما أجري من دراسات في هذا المجال قد كشف في الواقع عن العكس. وهناك عدد من العوامل التي تحدد المستوى الوظيفي المناسب للعاملين بالكشف، نذكر منها :

١ - مدى تعقد الموضوع المتداول .

٢ - نوعية لغة التكشيف المستخدمة؛ فمن الممكن للتكشيف المعتمد على الكلمات المفتاحية غير المقيدة أن يتطلب مكشفين أقل تمرساً ممن يمكن أن يتطلبهم استخدام لغة مقيدة ضخمة.

٣ - مدى الشمول والتخصيص في التكشيف ؛ فكلما ازدادت التفاصيل التقنية التي يتم تكشيفها ازدادت الحاجة إلى خبراء في الموضوع.

٤ - المرحلة التي يمر بها تطور النظام ؛ ففي المراحل المبكرة لتكشيف إحدى المجموعات بلغة مقيدة، عادة ما يكون كل قرار يتم اتخاذه بشأن التكشيف ، قراراً فكرياً. أما في المراحل التالية ، وحيث تكون القرارات الفكرية قد سجلت في ملف استنادي أو في لغة للدخول ، فإنه ربما أمكن لعملية التكشيف أن يعهد بها إلى فئة من العاملين ليسوا على مستوى عالٍ من التأهيل، باستطاعتهم النهوض بأكبر قدر من العمل باتباع القرارات التي سبق اتخاذها.

٥ - نوعية الأدوات المتوافرة للمساعدة في عملية التكشيف.

٦ - نوعية برنامج التدريب على التكشيف .

ومما لا شك فيه أنه لا يمكن لتحليل فعالية تكلفة النظام الفرعي للتكشيف أن يكتمل دون دراسة مستويات القوى البشرية اللازمة للعمل. ولتحقيق ذلك ينبغي تكشيف عدة مجموعات مختلفة من الوثائق من جانب فئات متفاوتة المستويات من القوى البشرية ، ثم مقارنة ناتج التكشيف بتكشيف معياري لوثائق الاختبار يتم إقراره باتفاق الآراء . وينبغي أخذ عوامل الوقت في الحسبان. ومن الممكن بعد ذلك مقارنة فعالية التكشيف التي حققتها هذه الفئات المختلفة ، والتي أمكن قياسها ، بتكاليف التكشيف.

استراتيجية البحث :

استراتيجية البحث ، كما بينا في الفصل الثامن ، بيان بشروط عضوية الوثائق بالفئات ، يتم التعبير عنه بالحاصل المنطقي أو الناتج المنطقي أو المتعمق المنطقي . وتستخدم أساليب تقاطع الفئات وجمعها وتكاملتها، للتنويع في مستويات الشمول والتخصيص في استراتيجيات البحث.

لننظر في طلب البحث الوارد في الشكل رقم (٥٨) والذي تم تحليله إلى أربعة جوانب

أو أوجه. ومن الممكن الربط بين المصطلحات الواردة في كل مجموعة من مجموعات المفاهيم أو الأوجه بواسطة « أو » البولينية ، كما أنه من الممكن الربط بين مجموعات المفاهيم بواسطة « و » البولينية .

| طلب بحث : العلاجات المستعملة للسجينات ذوات التاريخ في إساءة استعمال العقاقير استراتيجية البحث : | | | |
|--|---|---------------------------|--|
| أ | ب | ج | د |
| مفهوم الذات أو احترام الذات أو صورة الذات | المداداة أو العلاجات أو العلاج النفسي | السجناء أو السجينات | استخدام العقاقير أو سوء استخدام العقاقير |

الشكل رقم (٥٨) : طلب بحث واستراتيجية بحث (أسلوب قوالب البناء) .

وهناك طرق متعددة لتغيير مستوى الاستدعاء والتحقيق في هذا البحث ، وهذه الطرق نوجزها في الشكل رقم (٥٩) ، وتشمل الأساليب التي تستخدم خصائص اللغة الطبيعية واللغة المقيدة ، بالإضافة إلى أدوات البحث ، كالعوامل البولينية ، ومحددات الحقول ، وبناء العبارات (عوامل التقارب) والمترادفات ، والأشكال الهجائية المختلفة ، ورموز البتر.

ويؤدي استعمال « و » البولينية إلى زيادة الشمول في البحث ، ومن ثم زيادة التحقيق في الاسترجاع. وربما كان البحث الوارد في الشكل رقم (٥٨) شاملاً جداً في الواقع، وقد لا يؤدي إلى استرجاع أية وثائق. ومن الممكن الحد من الشمول بإغفال بعض الجوانب ، وربما بالبحث بدون الجانب « أ » (مفهوم الذات) أو بدون الجانب « ب » (العلاجات) . ويمكن للتأثير العام المترتب على ذلك أن يعزز الاستدعاء.

وهناك طريقة أخرى لتعديل هذا البحث ، وهي استعمال مصطلحات إضافية في كل مجموعة من مجموعات الأوجه أو المفاهيم المرتبطة ببعضها البعض بالعامل «أو» . وهذه يمكن أن تكون مصطلحات متصلة دلاليا (أوسع / أضيق ، أو متصلة مثل «سجين / سجينة) بالمصطلحات التي تم تعيينها ، أو مترادفات حقيقية ، أو أشباه مترادفات (مفهوم الذات / صورة الذات) ، أو أشكال هجائية مختلفة ، أو أسماء استهلاكية،

أو مصطلحات مبتورة ، يمكن أن تصل إلى جميع أشكال الكلمة ذات الجذر المحدد (THERAPY, THERAPIES, THERAPEUTIC) ويؤدي استعمال هذه الأدوات عموماً، للمضاهاة مقابل أعداد متزايدة من المصطلحات إلى تعزيز الاستدعاء ، أما الحد من استعمالها فيؤدي إلى ارتفاع نسبة التحقيق.

| لزيادة الاستدعاء | لزيادة التحقيق | أدوات البحث |
|--|---|----------------------------------|
| المصطلحات من اللغة الطبيعية أو من اللغة المقيدة أو منهما معا | | |
| استبعد مجموعات مفاهيم من الاستراتيجية (الربط بين عدد قليل من المفاهيم بواسطة «و» | أدخل مجموعات مفاهيم إضافية | العوامل البوليانية |
| أدخل مصطلحات إضافية في مجموعات المفاهيم (مرتبطة ببعضها بواسطة «أو» | استبعد مصطلحات من مجموعة المفاهيم المرتبطة ببعضها بواسطة «أو» | |
| المصطلحات من اللغة الطبيعية | | |
| استعمل مصطلحات إضافية عبارة عن أشكال هجائية مختلفة ، أو المختصرات ، أو أسماء استهلاكية | استعمل الأشكال الهجائية الشائعة فقط ، والمختصرات ، والأسماء الاستهلاكية | الهجاء والبتسر والمترادفات |
| ابتز حتى جذوع الكلمات | استعمل مصطلحات مخصصة دون بتر | |
| استعمل المترادفات | قلل من عدد المترادفات | |
| المصطلحات من اللغة الطبيعية | | |
| استعمل الكلمات المفردة أو العبارات القصيرة. | استعمل العبارات الطويلة | عوامل التقارب |

الشكل رقم (٥٩) : طرق زيادة الاستدعاء (يمكن أن تؤدي إلى نقص التحقيق) والتحقيق (يمكن أن يؤدي إلى نقص الاستدعاء) في بحث معين.

ولا ننسى أن الاستراتيجية التي تعرضنا لها حتى الآن تشتمل على مصطلحات قصد بها مضاهاة كل من اللغة الطبيعية واللغة المقيدة في الوثيقة. وينبغي مقارنة هذا الأسلوب، المؤلف الآن تماماً في البحث على الخط المباشر، بالحالة التي لا يستعمل فيها سوى واصفات اللغة المقيدة. وأوجه الاختلاف بين الحالتين بوجه عام هي:

١ - يتم في مرحلة المدخلات في اللغة المقيدة تقنين كثير من أدوات الاستدعاء التي عرضنا لها (كإدخال المترادفات، والأشكال الهجائية المختلفة، والبتة)، ومن ثم فإنه لا داعي لإدخالها بواسطة مسئول البحث الذي يقوم بإجراء بحث في لغة مقيدة بشكل صارم.

٢ - من الممكن استعمال الكلمات (التي تعزز الاستدعاء) أو العبارات المركبة (التي تعزز التحقيق) في بحث اللغة الطبيعية، إلا أن الصيغ المبسطة أو المركبة في هذا المستوى عادة ما يتحدد شكلها على نحو صارم في المكانز. والعملية المناظرة لذلك في اللغات المقيدة هي جمع الرؤوس الرئيسة والرؤوس الفرعية في تجميعات مؤتلفة مترابطة فيما بينها في مرحلة المدخلات.

٣ - من الممكن صياغة استراتيجيات البحث بحيث تستوعب وثائق مختلفة تتصل ببعضها البعض بعلاقات دلالية مختلفة، وذلك في مجموعة المفاهيم. ويمكن لبناء اللغة المقيدة أن يوحي بمثل هذه المصطلحات لمسئول البحث. وعلى مسئول البحث عند استعمال اللغة الطبيعية أن يتدبر كل هذه العلاقات، وأن يقدم المصطلحات اللازمة.

وعلى ذلك فإن أخطاء الاستدعاء في اللغة الطبيعية يمكن أن تقع عندما يعجز مسئول البحث عن تدبر مختلف المترادفات المحتملة، والأشكال المختلفة للكلمات، والمصطلحات الأخرى المرتبطة، الخاصة بمجموعة مفاهيم معينة. أما أخطاء التحقيق فيمكن أن تقع عندما يعجز مسئول البحث عن التحقق من جميع الأوجه التي يتضمنها البحث، أو عندما تسفر الكلمات المستعملة في مختلف التجميعات المنطقية المؤتلفة عن ارتباطات زائفة.

ويتضح من المناقشات السابقة أن هناك كثيراً من العوامل التي ينبغي أخذها في

الحسبان عند صياغة استراتيجية بحث ترمي إلى بلوغ أقصى مستويات التحقيق أو الاستدعاء أو كليهما معا. وقد ساد في الماضي الاعتقاد بأنه من الممكن للوسطاء المتمرسين، المتألفين فعلا مع هذه الأدوات وكيفية استعمالها، إجراء عمليات البحث للمستفيدين ، إلا أنه قد أمكن مؤخرا إدخال مخرجات وخدمات موجهة خصيصاً للمستفيد النهائي. وهذه الخدمات بوجه عام هي المنتجات نفسها التي كان يتم البحث فيها من قبل بواسطة الوسطاء، إلا أنها غالبا ما تقدم الآن مصحوبة بواجهة تعامل من البرمجيات التي تساعد المستجدين في صياغة استراتيجيات البحث.

وعلى ذلك ، فإنه يمكن لكم ما يتوافر لمسئول البحث من الأدوات المساعدة للاسترجاع على الخط المباشر، ونوعية هذه الأدوات، أن يؤثر في كفاءة الاسترجاع. وهناك نوعان مألوفان لواجهات التعامل مع نظم استرجاع المعلومات، وهما الواجهات التي تعمل بالأوامر **Command-driven** والواجهات التي تعمل بقوائم الاختيار **Menu-driven** ، وكلا النوعين من البرمجيات يمكن استخدامهما للبحث في الفهارس المتاحة على الخط المباشر وفي مرصد بيانات مقالات الدوريات. وفي دراسة تجريبية للمقارنة بين كفاءة فهرسين متوافرين على الخط المباشر، أحدهما يعمل بقوائم الاختيار والآخر يعمل بالأوامر ، تبين لجُـر ولسك (1983) **Geller and Lesk** أن المستفيدين يفضلون نظام الأوامر. وقد استخدم المستفيدون عمليات البحث بالكلمات المفتاحية اعتمادا على النظام الذي يعمل بالأوامر بكثافة أكثر من استخدامهم للنظام الذي يعمل بقوائم الاختيار ، وعثروا على أعداد كبيرة من الوثائق الصالحة. وقد درس كل من سوليفان وبورجمان وويبرن (1990) **Sullivan, Borgman and Wipperfurth** كلا من نظام الأوامر ونظام قوائم الاختيار في البحث في مرصدين للبيانات ، عن طريق دIALOG ، وتبين لهما أن استعمال واجهة قوائم الاختيار لم يؤثر في نوعية الاسترجاع أو مدى رضا المستفيد ، على الرغم من أن من تدربوا على استعمال الأوامر البسيطة قد استغرق تدريبهم وقتا أقل مما استغرقه التدريب على النظام الآخر ، كما أنهم تفاعلوا أيضا مع مرصدي البيانات أكثر مما تفاعل أقرانهم ممن تدربوا على نظام قوائم الاختيار.

هذا ، وتدل تقارير كفاءة البحث بواسطة المستفيدين النهائيين على أنه على الرغم من حماس كثير من المستجدين الواضح، لإجراء عمليات البحث الخاصة بهم بأنفسهم،

فإن عمليات البحث هذه ليست على مستوى عال من الفعالية، خاصة في حالة عدم تدريب المستخدمين على التعامل مع النظام، أو في حالة عدم التعامل مع النظام بكثافة (انظر على سبيل المثال Lancaster et al. 1992). وتدل التحليلات التفصيلية لسجلات الواقعات (سجلات إلكترونية لجميع عمليات البحث التي يتم إجراؤها في نظام ما خلال فترة بعينها) على أن المستخدمين النهائيين يميلون لاستعمال مصطلحات بحث أقل مما يستعمله الوسطاء، مما يترتب عليه أخطاء في التحقيق (نتيجة لقصور الشمول في الاستراتيجية) وأخطاء في الاستدعاء (نتيجة لقلة أعداد المترادفات والمصطلحات المرتبطة التي تشتمل عليها مجموعات المفاهيم). هذا بالإضافة إلى أن المتمرسين من الوسطاء يسترجعون، في حدود زمنية معينة، تسجيلات مناسبة أكثر من تلك التي يسترجعها الباحثون غير المتمرسين (Penhale and Taylor 1986).

وللبحث في الفهارس المتاحة على الخط المباشر مشكلاته الخاصة، فانخفاض مستوى الشمول في الكشف باستعمال قائمة رءوس موضوعات مكتبة الكونجرس *Library of Congress Subject Headings* مقرونا بالربط المسبق المخصص جدا للرءوس الرئيسية والرءوس الفرعية، يجعل من الصعب البحث في مثل هذه الفهارس. وللحصول على استدعاء مقبول يتعين إجراء عمليات بحث عريضة جدا، غالباً ما تسفر عن استرجاع المئات من الوثائق غير الصالحة، وبذلك ينخفض التحقيق بشكل ملحوظ (Lancaster et al. 1990)، وعند إجراء عمليات بحث أكثر تخصيصاً ينخفض الاستدعاء بشكل ملحوظ، كما أن المصطلحات التي يدخلها المستخدمون قلما تضاهاي أيًا من رءوس الموضوعات (Vizine-Goetz and Drabenstott 1991).

وهناك عملية أخرى يمكن أن يجريها من يقوم بصياغة استراتيجية البحث، ويمكن أن يكون لها بعض الأثر على الاستدعاء والتحقيق في نتائج البحث، وهي على وجه التحديد عملية تنقية أو غربلة المخرجات؛ ففي بعض مرافق المعلومات يقوم مسئول البحث بفحص مخرجات النظام بعناية قبل تقديم النتائج للمستخدم، في محاولة لاستبعاد التسجيلات التي تبدو بوضوح غير صالحة، وبذلك يحسن نسبة التحقيق في الناتج النهائي الذي يتلقاه المستخدم. ويرتبط نجاح عملية التنقية هذه مباشرة بنوعية بيان الطلب، نظراً لأنه بناء على هذا الأساس يتم اتخاذ قرارات الصلاحية من جانب

مسئول البحث. فإذا كان بيان الطلب تعبيراً غير مكتمل عن حاجة المستفيد من المعلومات، فإنه يمكن لمحلل الطلب أن يستبعد تسجيلات كان من الممكن للمستفيد أن يحكم بصلاحياتها كما يستبعد تسجيلات كان من الممكن الحكم بعدم صلاحيتها.

ومن العوامل الأخرى المؤثرة في عملية التنقية مقدار الوقت المستنفد في العملية، ونوعية بدائل الوثائق التي يقدمها النظام؛ فكلما كانت البدائل مكتملة كان من الأسر على مسئول البحث إصدار أحكام الصلاحية بدقة. فمن الممكن للعناوين مضافاً إليها المصطلحات الكشفية أن تكون أفضل دلالة من العناوين وحدها، كما يمكن للعناوين مضافاً إليها المستخلصات أن تكون أفضل دلالة من العناوين مضافاً إليها المصطلحات الكشفية. ومن الممكن في الواقع للاعتماد على تسجيلات الوثائق في إصدار الأحكام حول صلاحية الوثائق بالنسبة لبيان الطلب، أن يكون مرتبطاً وبشكل مباشر بطول التسجيلية.

جوانب فعالية التكلفة في عمليات البحث:

من الممكن أن تكون هناك علاقة تأثير وتأثر بين الجهد المبذول في صياغة استراتيجية البحث والجهد المستنفد في تنقية مخرجات البحث؛ فإذا أنفق مسئول البحث قدراً كبيراً من الوقت في صياغة استراتيجية بحث دقيقة محكمة البنيان، فإنه من الممكن الحصول على مخرجات بحث مرتفعة التحقيق، تتطلب قدراً ضئيلاً من التنقية أو التحرير، وقد لا تتطلب ذلك مطلقاً. أما إذا تم إجراء بحث عريض، فسوف يتطلب الأمر بذل مزيد من الجهد في التخلص من التسجيلات التي تبدو بوضوح غير صالحة، في أثناء عملية التنقية. وإذا كان بإمكان مسئول البحث النظر في المخرجات واتخاذ قرارات الصلاحية التي تضاهي وبشكل معقول تلك التي كان من الممكن اتخاذها من جانب المستفيد، فإنه من الممكن لهذا الأسلوب الأخير أن يكون أكثر فعالية من حيث إحراز استدعاء مرتفع في مقابل تحقيق يمكن للمستفيد تحمله. كذلك يمكن لهذا الأسلوب أن يكون أكثر فعالية من وجهة نظر التكلفة.

ولدى التفاعل الذي يتم بين محلل الطلب والمستفيد انعكاساته أيضاً على فعالية التكلفة. فالأمر يستلزم قدراً من التفاعل لكي يكون بيان الطلب معبراً بدقة عن حاجة المستفيد من المعلومات. وكلما ازداد التفاعل بين الطرفين ازدادت احتمالات تحسن

الاستدعاء أو التحقيق. إلا أنه من الممكن أن يكون هناك أيضا احتمالات مواءمة ، بين مقدار الجهد المبذول في التفاعل مع المستفيد قبل إجراء البحث والجهد المستنفد في تنقية المخرجات ، مثلا. أضف إلى ذلك ، أن هناك عدة مراحل يمكن أن يتم فيها التفاعل ؛ مرحلة تقديم الطلب ، ومرحلة صياغة استراتيجية البحث ، ومرحلة تقديم الاستراتيجية المقترحة للمستفيد لإقرارها أو تعديلها ، ومرحلة المخرجات ، أي مرحلة إجراء البحث بشكل تكراري ، حيث يتم تقييم النتائج الأولية للبحث من جانب المستفيد ، وصياغة استراتيجية جديدة ، يدويا أو آليا ، بناء على تقييم صلاحية التسجيلات المسترجعة . وكلما تأخر حدوث التفاعل في عملية الاسترجاع كان ذلك أفضل ؛ فقد تبين للسك وسالتون (Lesk and Salton 1969) في تقييمهما لعمليات البحث في نظام سمارت SMART، أن التفاعل بعد إجراء البحث عادة ما يكون أفضل من التفاعل قبل عملية البحث . وهناك بوجه عام ، مجموعة مثل من إجراءات التفاعل في بيئة معينة ، كما أن هناك بعض طرق التفاعل التي تؤدي فعلا إلى تدني كفاءة البحث لا إلى ارتفاعها . فقد تبين على سبيل المثال ، في تقييم المدلرز (Lancaster, 1968a) أن المقابلات المباشرة ، وجها لوجه بين المستفيدين ومسئولي إجراء عمليات البحث ، أو بين المستفيدين واختصاصيي المكتبات الطبية ، لم تكن تؤدي بالضرورة إلى تحسن نوعية الطلبات ، كما أدت في بعض الحالات إلى تشويه الطلب أو ضياع معالمة . ويبدو أنه من الأفضل من وجهة نظر الفعالية أن يطلب من المستفيد أن يقدم الطلب مكتوبا في البداية ، بمصطلحات لغته الطبيعية ، ثم التفاعل بعد ذلك لاستيضاح بيان الطلب ، بدلا من أن يناقش المستفيد حاجته من المعلومات مع اختصاصي المعلومات ، الذي يقوم بعد ذلك بإعداد بيان الطلب .

ويمكن للتفاعل بين المستفيد والنظام أن يكون بالنسبة لفعالية التكلفة ، أهم مكونات نظام الاسترجاع على الإطلاق. فإذا تقدم المستفيد بطلب قاصر ، لا يعبر بدقة عن الحاجة إلى المعلومات ، فإن مصير عملية البحث هو الفشل المؤكد تقريبا ، مهما بلغت كفاءة الكشف ، ولغة النظام ، واستراتيجيات البحث ، وبذلك يضيع جهد البحث هباء . ولتجنب تبديد الوقت والجهد بلا طائل ، في صياغة استراتيجيات البحث ، وفي تجهيز عملية البحث، وفي تنقية المخرجات ، فإن الأساليب الكفيلة بتحسين نوعية الطلبات هي الأيسر عادة في تبريرها على أسس اقتصادية. فمن الممكن على سبيل المثال ، أن يتحقق

قدر كبير من التحسن نتيجة لتوافر نموذج محكم البنيان ، مصمم لمساعدة المستفيد على تقديم أفضل طلب ممكن للنظام . وربما تتطلب تعبئة النموذج مزيداً من الجهد من جانب المستفيد في البداية ، إلا أنه يمكن أن يؤدي إلى تحسن ناتج البحث ، كما يؤدي إلى الاقتصاد في الوقت على المدى الطويل ، وخاصة الوقت المستنفد في التنقية من جانب المستفيد أو من جانب المسئول عن إجراء عملية البحث.

ولقد تبين وبشكل قاطع (على يدي فيتزجيرالد Fitzgerald, 1981 على سبيل المثال) أن التفاعل بين المستفيد واختصاصي المعلومات أثناء إجراء عملية البحث ، أي الجلوس معاً إلى المنفذ ، عادة ما يسفر عن أفضل نتائج ، على الرغم من أنه باهظ التكلفة ولاشك ، نظراً لأنه ينطوي على استنفاد وقت كل من الطرفين.

عمليات المواءمة في استرجاع المعلومات :

هناك كثير من حالات المواءمة المحتملة بين مختلف عمليات استرجاع المعلومات ، فقد تعرض القسم السابق ، على سبيل المثال ، للجهد المستنفد في الكشف واللغة في مقابل الجهد المستنفد في إجراء عملية البحث ، والجهد المستنفد في صياغة استراتيجية البحث في مقابل الجهد المستنفد في تنقية المخرجات. ويقارن تحليل فعالية تكلفة النظام بأكمله بين المواءمات المحتملة ، وبين أكثر التجمعات المؤتلفة من الإجراءات كفاءة ، لبلوغ مستوى أداء معين يتناسب مع التكلفة.

وأهم مواءمة ينبغي أخذها في الحسبان ، هي المواءمة الشاملة بين تكاليف المدخلات وتكاليف المخرجات ، ودائماً ما تسفر تدابير الاقتصاد في تكلفة إجراءات المدخلات ، عن تزايد العبء على عمليات المخرجات ، وبذلك تزداد تكاليف المخرجات . وفي مقابل ذلك ، فإن الحرص الزائد في تجهيز المدخلات ، والذي يعني عادة تزايد تكاليف المدخلات ، يمكن أن يؤدي إلى ارتفاع مستوى الأداء ، والحد من تكاليف المخرجات . وتشتمل القائمة التالية على حصر لبعض المواءمات المحتملة:

١ - لغة الكشف المقيدة محكمة البنيان في مقابل الاستعمال الحر للكلمات المفتاحية غير المقيدة : تستنفد اللغة المقيدة جهداً في بنائها وصيانتها ، كما أنها أكثر تكلفة في استخدامها في الكشف ؛ فاختيار المصطلحات من اللغة المقيدة ، والذي يمكن أن

ينطوي على عملية مراجعة بعض الأدوات المساعدة ، يستنفد بوجه عام ، وقتاً أطول مما يستنفده تعيين الكلمات المفتاحية غير المقيدة . هذا بالإضافة إلى أن التكشيف بالكلمات المفتاحية يمكن أن يتطلب عناصر بشرية بمؤهلات أدنى من تلك التي يتطلبها استخدام اللغة المقيدة الأكثر تعقداً. إلا أن اللغة المقيدة تؤدي إلى الاقتصاد في الوقت والجهد في مرحلة المخرجات ، ويفرض البحث في اللغة الطبيعية أو بالكلمات المفتاحية دون الاستعانة بلغة مقيدة ، عبئاً إضافياً على مسئول البحث الذي يضطر تقريباً لبناء قطاع من لغة مقيدة في كل مرة يعد فيها لإجراء عملية بحث ، وذلك بأن يتدبر ، على سبيل المثال جميع الطرق المحتملة للتعبير عن «الكيمائيات النفطية» أو «صناعة النسيج» بالكلمات المفتاحية أو بنصوص اللغة الطبيعية . كذلك يمكن للاستعمال غير المقيد للكلمات المفتاحية أن يؤدي إلى انخفاض متوسط التحقيق في البحث ، وبذلك يتطلب جهداً إضافياً وتكلفة في تنقية المخرجات.

٢ - الضبط الصارم للجودة في التكشيف عن طريق عملية المراجعة مثلاً ، في مقابل التكشيف دون إجراء للمراجعة : تؤدي المراجعة إلى زيادة تكاليف التكشيف ، إلا أنها من المفترض أن تؤدي إلى الاقتصاد في تكاليف المخرجات وذلك بالحد من الوقت المستنفد في التنقية اللازمة للتخلص من الوثائق التي يتبدى عدم صلاحيتها بشكسل واضح. وما إذا كان لمراجعة المدخلات ما يبررها اقتصادياً ، أمر لا يمكن حسمه إلا بتقييم عدد الأخطاء التي تقع في التكشيف ، وعدد الأخطاء التي يمكن تصويبها بإجراءات ضبط الجودة.

٣ - اللغة المقيدة عالية التخصيص في مقابل اللغة المقيدة العريضة : فالأولى أعلى تكلفة في وضعها وصيانتها وتطبيقها. وكلما ازدادت اللغة تخصيصاً ازدادت صعوبة تحقيق الاطراد في التكشيف ، وارتفع مستوى القوى البشرية التي يمكن أن تدعو الحاجة إليها لاستعمال هذه اللغة. ومع ذلك ، فإن اللغة عالية التخصيص يمكن أن تكفل الارتفاع بنسبة التحقيق في البحث، وبذلك تؤدي إلى الاقتصاد في الوقت المستنفد في تنقية المخرجات. وهناك شكل معين من التخصيص يتحقق بواسطة مؤشرات الدور أو الروابط ، وتنطبق هذه الملاحظات أيضاً على استعمال مثل هذه الأدوات.

هذه مجرد أمثلة ثلاثة للمواءمات المحتملة بين ما يبذل من جهد في المدخلات وما يبذل في المخرجات . وهناك الكثير من الاحتمالات الأخرى . ويقدم الشكل رقم (٦٠) مقارنة للمواءمة في نظامين افتراضيين لاسترجاع المعلومات . ففي النظام « أ » تحظى عمليات المدخلات بقدر كبير من العناية والتكاليف ، مما يؤدي إلى الاقتصاد في الجهد والتكاليف في المخرجات . أما في النظام « ب » فقد تم الاقتصاد في تكاليف المدخلات ، ومن ثم تزايد الجهد والتكاليف في المخرجات . وليس من الضروري أن يكون أحد النظامين أكثر كفاءة من الآخر . وربما كان الأسلوب المتبع في النظام « ب » أكثر فعالية من وجهة نظر التكلفة من الأسلوب المتبع في النظام « أ » ، إذا حقق مستوى أداء مقبولا بالنسبة للمستفيد النهائي ، بتكلفة إجمالية أقل من تلك التي أنفقت على النظام « أ » .

وهناك عدد كبير من العوامل المختلفة المؤثرة في القرار الخاص بالتركيز على عمليات المدخلات أو على عمليات المخرجات في نظم استرجاع المعلومات . وربما كانت أهم الاعتبارات ما يلي :

١ - عدد الوثائق التي يتم تكشيفها وعدد الطلبات التي يتم تجهيزها سنوياً : ففي الموقف المتطرف ، حيث يتم تكشيف أعداد كبيرة من الوثائق وتجهيز عدد ضئيل من الطلبات ، ربما كان من الأوفق ، إذا ما تساوت جميع الأمور الأخرى ، الاقتصاد في تكاليف المدخلات وإلقاء أعباء إضافية على مرحلة المخرجات . أما في الموقف العكسي ، حيث يتم إدخال عدد قليل من الوثائق في مقابل تجهيز أعداد كبيرة من الطلبات ، فإن العكس قد يكون هو الصواب .

٢ - السرعة المطلوبة في المدخلات : يتحتم في بعض المواقف إدخال الوثائق في النظام بأقصى سرعة ممكنة ، ويصدق ذلك بوجه خاص حيثما يكون النظام في خدمة وظيفة البحث (الإحاطة الجارية) . ومن الممكن في ظل هذه الظروف ، للسرعة اللازمة في المدخلات أن تفوق كل ما عداها من اعتبارات ، بالإضافة إلى اتخاذ تدابير الاقتصاد في التكاليف ، لإدخال التسجيلات في مرصد البيانات بأقصى سرعة ممكنة .

٣ - السرعة المطلوبة في المخرجات : من الممكن في مواقف أخرى أن تكون الاستجابة السريعة الدقيقة أمراً حيوياً ، كما هو الحال مثلاً في مركز معلومات السموم ، كما أن

| النظام (ب) | النظام (أ) |
|---|--|
| <p style="text-align: center;">خصائص المدخلات</p> <p>لغة مفيدة صغيرة يكملها الاستعمال الحر لل كلمات المفتاحية تكشيف منخفض الشمول (■ مصطلحات للوثيقة) مكشوفون أقل في مستوى التأهيل ، من غير الحاصلين على مؤهلات جامعية . لا مراجعة للتكشيف متوسط إنتاجية المكشف من ١٠٠ - ١٥٠ وثيقة في اليوم تكاليف مدخلات منخفضة بث سريع</p> | <p>لغة ضخمة محكمة التقنين تكشيف متوسط الشمول (١٠ مصطلحات للوثيقة في المتوسط) مكشوفون مؤهلون تأهيلاً عالياً يتقاضون مرتبات مرتفعة مراجعة التكشيف متوسط إنتاجية المكشف ١٠٠ وثيقة في اليوم تكاليف مدخلات مرتفعة فاصل زمني طويل نسبياً بين نشر الوثائق وإدخال بياناتها في النظام.</p> |
| خصائص المخرجات | |
| <p>عبء زائد على مسئول البحث في إعداد الاستراتيجيات تحقيق منخفض في الناتج الخام استدعاء يمكن تحمله الحاجة إلى تنقية الناتج الخام للارتفاع بالتحقيق إلى مستوى يمكن تحمله من جانب المستفيد النهائي. استجابة بطيئة تكاليف بحث مرتفعة نسبياً</p> | <p>عبء مخفف على مسئول البحث في إعداد الاستراتيجيات تحقيق مرتفع في الناتج الخام استدعاء يمكن تحمله لا حاجة إلى التنقية زمن استجابة سريع تكاليف بحث منخفضة نسبياً</p> |

الشكل رقم (٦٠): مقارنة بين المواءمة في نظامين افتراضيين لاسترجاع المعلومات .

تدابير الاقتصاد في المدخلات لاتجد لها مبرراً إذا كانت تؤدي إلى تأخر الاستجابة أو انخفاض مستوى الدقة في المخرجات.

٤ - المنتجات الجانبية : قد يكون من الممكن في بعض الأحيان الحصول على مرصد بيانات قابل للبحث بتكلفة منخفضة جداً ، كأن يكون على سبيل المثال مرصداً إلكترونياً للبيانات ، يعتمد على اللغة الطبيعية ، عبارة عن ناتج جانبي لبعض العمليات الأخرى كالنشر أو إعداد التقارير مثلاً ، أو يكون هذا المرصد متاحاً من جانب مركز معلومات آخر . وعلى الرغم من أن صيغ المدخلات ونوعية هذه المدخلات قد لاتكون مثالية ، فإن مرصد البيانات إذا ما كان متاحاً بتكلفة رمزية ، فإنه قد يكون مرغوباً ، من وجهة نظر فعالية التكلفة ، للإفادة منه ، وربما مع بعض التعديلات الطفيفة ، مع الاستعداد لبذل مزيد من الجهد في عملية البحث.

هذا ، وقد سبق أن تعرضنا لبعض عوامل فعالية التكلفة المتصلة بمختلف النظم الفرعية لنظام استرجاع المعلومات المكتمل ، وهي التكشيف ، ولغة التكشيف ، والبحث ، وتفاعل المستفيد مع النظام ، وكما هو الحال في تقييم الفعالية ، فإنه مما يتجاوز الحدود الواقعية ويعد من قبيل المخاطرة ، في تحليل فعالية التكلفة ، النظر في أي من هذه النظم الفرعية بمعزل عن السياق العام . فكل هذه المكونات ترتبط ببعضها البعض ارتباطاً وثيقاً ، ودائماً ما يكون لأي تغير جوهري في أحدها أن يكون له حتماً انعكاساته على النظام برمته . وينبغي أن يكون مسئولو التقييم على دراية بذلك ، وأن يأخذوا في حساباتهم ، في أي تحليل لفعالية التكلفة ، الآثار غير المباشرة ، بعيدة المدى للتغيرات التي تطرأ على النظام ، فضلاً عن الآثار المباشرة الفورية . لنفترض على سبيل المثال ، أنه قد اتخذ قرار للتحويل من لغة تكشيف مقيدة ، متطورة محكمة البنيان ، إلى شيء آخر أبسط ، حينئذ يمكن للآثار الفورية أن تتمثل في الحد من تكاليف التحكم في اللغة وصيانة هذه اللغة ، والاقتصاد في وقت التكشيف ، والتحسين في زمن الاستجابة . هذا بالإضافة إلى بعض الآثار غير المباشرة إلى حد ما ، بعيدة المدى ؛ فمن الممكن للوقت اللازم لإعداد استراتيجيات البحث أن يزداد ، مما يؤدي إلى زيادة تكاليف البحث ، كذلك يمكن لنسبة التحقيق أن تنخفض ، بحيث يمكن أن تدعو الحاجة إلى وجود تدابير خاصة بتنقية المخرجات ، وإذا دعت الحاجة إلى تنقية المخرجات ، فإن نوعية بدائل

الوثائق في النظام قد تحتاج إلى تطوير ؛ وربما تطلب الأمر إدخال المستخلصات ، على الرغم من أنها لم تكن ضرورية من قبل .

كذلك يمكن أن تحدث ظاهرة مماثلة إذا ما ارتفع مستوى الشمول في الكشف ؛ حيث يمكن أن تحدث زيادة فورية في زمن الكشف وفي تكلفته ، وزيادة أيضا في متوسط عدد الوثائق التي تسترجع في البحث ، فضلاً عن التحسن في الاستدعاء ، والانخفاض في متوسط التحقيق. أما على المدى البعيد ، فإن الأمر قد يتطلب اتخاذ تدابير تنقية المخرجات وتحسين نوعية بدائل الوثائق ، وذلك لضمان بقاء التحقيق في مستوى يمكن تحمله. ومما لا شك فيه أن نظام المعلومات كائن مركب ، ولما يطرأ عليه من تغيرات أثارها بعيدة المدى.

العوامل البشرية في الكشف والبحث :

يصور الشكل رقم (٦١) الخطوات التي تنطوي عليها عملية الاسترجاع ، والتي ترتبط بقابلية النظام للتغير . وعلى ذلك ، فإن محتويات مرصد البيانات ، وسياسات الكشف ، ولغة الاسترجاع ، وعمليات البحث ، كل هذه يمكن أن تتغير وبطرق شتى ، وبما يترتب على ذلك من آثار على الأداء سبق أن تعرضنا لها . إلا أن هناك بعض العوامل الأخرى ، المرتبطة ارتباطاً وثيقاً بإنشاء مرصد البيانات والإفادة منها ، والتي تؤثر في نظم الاسترجاع بشكل نسقي . وتشمل هذه العوامل مختلف خصائص المكشفين ومسئولي إجراء عمليات البحث ، ومن أبرزها الإلمام بمجال موضوعي معين والخبرة بنظام الاسترجاع نفسه .

ويتناول البحث العلمي في هذا المجال الفروق الفردية بين المكشفين ومسئولي عمليات البحث ، ويؤكد ساراسفك (1991) Saracevic أهمية هذا البحث وجدواه على النحو التالي :

« للبحث في الفروق الفردية في استرجاع المعلومات أهميته لعدة أسباب ، وها هي ثلاثة منها : أولها أن مثل هذا البحث يسهم في زيادة إدراكنا لأوجه الاختلاف في الأداء البشري في مهام استرجاع المعلومات ، تلك الفروق التي تلعب دوراً جوهرياً (إن لم يكن الدور الجوهري) في الأداء العام لأي نظام ولجميع نظم استرجاع المعلومات . أما السبب الثاني ، فإن مثل هذا البحث ضروري للتنبؤ في النهاية بفروق الأداء من جانب البشر ذوي الخصائص المختلفة ، حيث يمكن الاستفادة من هذه الفروق في تصميم نظم استرجاع المعلومات ، وتطوير إمكانات واجهات التعامل لكي تكون قادرة على الاستجابة

للشخص والتكيف معهم (سواء كانوا من الوسطاء أو من المستفيدين النهائيين) على اختلاف خصائصهم، بدلا مما هو عليه الحال الآن، حيث يعامل جميع المستفيدين على قدم المساواة. ويعني ذلك أنه من الممكن لمثل هذا البحث أن يفيد النظم في البدء في إدخال بعض مقومات فهم البشر. أما السبب الثالث، فقد أدت التطورات التقنية الراهنة، في كل من العتاد والبرمجيات إلى زيادة إمكانات المرونة، بحيث ازدادت احتمالات استيعاب ما بين المستفيدين من اختلافات. ومن هنا تتزايد صلاحية نتائج البحوث بالنسبة للتطورات التقنية» (P.82)

أضف إلى ذلك أن هناك نوعين أساسيين من الأسئلة التي يمكن للبحث في الفروق الفردية أن يسهم في الإجابة عنها :

« ١ - [الأسئلة المتعلقة بالاطراد.] إلى أي حد يسفر أداء البشر المتمتعين ببعض الخصائص المشتركة (من المكشفين ومستولي البحث مثلا) عن نتائج متشابهة عند القيام بنفس المهام في نظم استرجاع المعلومات (أي تكشف نفس المجموعة من الوثائق، أو البحث استجابة لنفس الطلب مثلا) ؟ أو بعبارة أخرى : ما مدى ضخامة أوجه الاختلاف أو الفروق ؟

« ٢ - [الأسئلة المتعلقة بكفاءة الأداء.] أي إلى أي مدى تؤثر مظاهر الاختلاف في خصيصة أو خصائص معينة (كمستوى الخبرة مثلا في عمل كالتكشيف أو إجراء عمليات البحث) في كفاءة الأداء والنتائج ؟ (Saracevic, 1991, P. 82)

دراسات الاطراد :

تقيس دراسة الاطراد مدى اتفاق فردين أو أكثر فيما بينهم ، ومدى اطراد أو اتساق شخص مع نفسه في بعض الأحكام أو المهام . وقد أجريت مثل هذه الدراسات في تعيين المصطلحات الكشفية ، وفي ترجمة الاستفسارات إلى استراتيجيات للبحث وفي أحكام صلاحية الوثائق المسترجعة.

وقد تم تجريب أو اقتراح عدد من المقاييس المحتملة للاطراد ، ومن بين المقاييس المألوفة «زوجية الاطراد (CP) Consisteney Pair» . ومن الممكن حساب زوجية الاطراد الخاصة باثنين من المكشفين «أ» و «ب» على النحو التالي :

عدد المصطلحات المعينة من قبل كل من أ و ب

عدد المصطلحات المعينة من قبل أ أو ب أو كليهما

وعلى ذلك ، فإنه إذا عين المكشف «أ» المصطلحات أ، ب، ج، د، هـ، و، لوثيقة

ما وعين المكشف « ب » المصطلحات أ ، ب ، ج ، ز ، ح ، للوثيقة نفسها، فإن قياس
الاطراد يتم على النحو التالي :

$$\frac{\text{أ، ب، ج}}{\text{أ، ب، ج، د، هـ، و- ز، ح}} \text{ أي } \frac{3}{8} \text{ أي } 37,5\%$$

ومن الممكن استخدام هذا المقياس نفسه للتعبير عن الاطراد أو الاتفاق بين
المكشفين المختلفين (الاتفاق بين المكشفين) أو بين القرارات التي يتخذها المكشف نفسه
في أوقات مختلفة (اطراد قرارات المكشف).

وقد بذلت جهود جادة لدراسة الاطراد في ستينيات القرن العشرين وسبعينياته ،
وحظيت هذه الجهود بالتغطية المكثفة من جانب أوليفر وآخرين (Oliver et al. (1966 ،
وكوبر (1984) Cooper ، وليونارد (1975) Leonard ، وماركي (1984a) Markey ، كما
تناولها مؤخراً أيضاً سيفرت وأندروز (1991) Sievert and Andrews . والنتيجة العامة
التي انتهت إليها هذه الدراسات ، هي أن أرقام الاتفاق أو الاطراد تميل للانخفاض
الشديد. وتمثل النتائج التي انتهى إليها زوند ودكستر Zunde and Dexter من أن
الاتفاق بين زوجيات المكشفين يتراوح في المتوسط بين ٢٤ بالمئة و ٤١ بالمئة ، نتائج هذه
الدراسات خير تمثيل. وقد تبين من عدد من الدراسات أن الاطراد في الكشف يتأثر
بعدة عوامل ، تشمل:

١ - الشمول في الكشف : فكلما ارتفع مستوى الشمول في الكشف ازدادت احتمالات
عدم الاطراد.

٢ - نوعية اللغة المستعملة : ويشمل ذلك حجم اللغة ومدى التخصيص فيها . وكلما
زاد حجم اللغة وازدادت تعقدا ، ازدادت احتمالات عدم الاطراد.

٣ - نوعية الأدوات المتاحة للمساعدة في الكشف : فكلما قلت الأدوات المساعدة المتاحة
ازدادت احتمالات عدم الاطراد.

كذلك يتأثر الاطراد في الكشف بالعوامل البشرية ، وفي مقدمة هذه العوامل ،
عاملان أولهما خبرة المكشفين وما يتلقون من تدريب ، وثانيهما إلمام المكشفين بالمجال

الموضوعي لرصد البيانات . ويلخص ساراسفك (1991) Saracevic النتائج المتاحة في الإنتاج الفكري على النحو التالي :

«كان متوسط الاطراد ١٠٪ عند اكتشاف براءات الاختراع الكيميائية بواسطة المكشفين المتمرسين والمكشفين غير المتمرسين.

كانت نسبة الاطراد تتراوح بين ٣٥٪ و ٤٥٪ بالنسبة للمكشفين المتمرسين الذين توافرت لهم أدوات مساعدة كاللغة المقيدة مثلاً، و ١٦٪ بالنسبة للمتمرسين و ١٢٪ بالنسبة لغير المتمرسين في حالة عدم توافر أدوات مساعدة»

تراوحت نسبة الاطراد بين ٢٦٪ و ٥٩٪ عندما قام غير المتخصصين في علم النفس، والمتخصصون في علم النفس ، بتكشيف المستخلصات النفسية . (P.83)

كذلك تمت دراسة الاطراد في قرارات الصلاحية في العقدين السابع والثامن من القرن العشرين، وقد حظيت هذه الدراسات بالمراجعة من جانب (1990) Nilon و (1975) Saracevic و Schomber، وتمثل النتائج التي انتهت إليها التجربة الميدانية الضخمة التي أجرتها جامعة كيس وسترن ريزيرف (1967) Rees and Schultz في مجال الطب، والتي شملت ١٨٤ مبحثاً على درجات متفاوتة من الخبرة في عمليات استرجاع المعلومات ، والمعرفة التخصصية في الطب تمثل نتائج هذه التجربة ، نتائج هذه الدراسات خير تمثيل ؛ فمن بين ما تبين للدارسين أن أحكام الصلاحية كانت أكثر اطرادا بين مسؤولي الاسترجاع المتمتعين بالمعرفة التخصصية أكثر من غيرهم ، وأن أحكام الصلاحية كانت مطردة بين خبراء الطب وبين المكتبيين العاملين بالمكتبات الطبية بنسبة ٥٠٪ تقريبا ، إلا أن المكتبيين العاملين بالمكتبات الطبية كانوا يميلون لأن يكونوا أكثر تساهلا من خبراء الطب في استعمالهم لميزان الصلاحية.

وقد قام كل من بيتس (1977) Bates وساراسفك وكانكتور Saracevic and Kantor (1988a,b) وساراسفك وآخرون (1988) Saracevic et al. بدراسة الاطراد في اختيار مصطلحات الاستفسارات من جانب مسؤولي البحث والاسترجاع . وقد تبين لبيتس انخفاض نسبي في الاطراد في اختيار رؤوس موضوعات مكتبة الكونجرس ، بين مسؤولي البحث في الفهارس المتاحة على الخط المباشر ، في مجالات علم النفس والاقتصاد وعلم المكتبات. وكانت نتائجها تتفق وما انتهى إليه كل من ساراسفك وكانكتور ، وساراسفك وآخرون ، اعتمادا على أربعين طلباً تم البحث عن الوثائق التي يمكن أن تلبيها ، في نظام ضخم لاسترجاع المعلومات ، جاهز للعمل (ديالوج). وقد تلقى مجموعة من مسؤولي

البحث المحترفين الاستفسار نفسه، حيث تم حساب زوجية الاطراد بالنسبة لجميع زوجيات مسئولى البحث المحتملة. وكان متوسط الاتفاق في اختيار مصطلحات عمليات البحث بالنسبة للأسئلة نفسها ٢٧ بالمئة.

ومن الممكن من دراسات الاطراد ، استخلاص عدد من النتائج العامة ؛ فتدل أولا النتائج الإجمالية المستخلصة من الدراسات التي أجريت في المجالات الثلاثة ، وهي الكشف وأحكام الصلاحية ، واختيار مصطلحات عمليات البحث ، على انخفاض الاطراد.^(١) إلا أنه على الرغم من أن هذه الدراسات تكشف عن عدم الاطراد ، فإنه من الممكن لعدم الاطراد أن يكون جيدا وسيئا في الوقت نفسه؛ فمن الممكن على سبيل المثال أن يتبين من إحدى الدراسات أنه من بين مجموعة من المكشفين قوامها ستة، هناك مكشfan أكثر اتساقا فيما بينهما من أي ثنائي آخر من المكشفين في المجموعة . وهناك مكشف واحد هو الأكثر تقردا بين الجميع ، وهو الأقل اتفاقا في المتوسط عند مقارنة أدائه في الكشف بأداء غيره من المكشفين في المجموعة. فليس من الضروري أن يكون أداء الفردين اللذين يسجلان أقصى درجات الاتفاق بينهما في الكشف هو أفضل أداء . وربما يتبين في الواقع أن المكشف المتفرد هو الأكثر فعالية لأن مصطلحاته تضاهي ما يقدم للنظام من استفسارات على أحسن وجه ، أي أن تكشيفه يكفل استرجاع أعداد كبيرة من الوثائق التي يحكم المستفيدون بصلاحياتها ، ويحول دون استرجاع كثير من الوثائق غير الصالحة.

دراسات الكفاءة :

على الرغم من تمتع دراسات الاطراد بقدر من الأهمية ، فإنه لا يمكن الخروج بنتائج عامة حول كفاءة الأداء إلا بدراسة ما بين العوامل المختلفة ومقاييس الفعالية من علاقات. وقد كرسنا الجانب الأكبر في هذا الفصل لأوجه النظام وكيف تؤثر في الكفاءة كما تقاس بالاستدعاء والتحقيق . ومن الممكن أيضا دراسة كيفية تأثير العوامل البشرية في هذه المقاييس.

(١) إلا أن فكرة الانخفاض هنا ، كما أشار ساراسفك وكانتور (Saracevic and Kantor (1988b) «قد لا تكون سليمة؛

فنحن لا نعرف ما هو " المرتفع " وربما كانت المعدلات التي تمت ملاحظتها هي أقصى ما يمكن توقعه ، أي أنها

يمكن أن تكون "طبيعية" . (P.207)

تركزت دراسات الفروق الفردية في استرجاع المعلومات على خصائص مسئول البحث ، وخاصة الخبرة في إجراء عمليات البحث ، والمعرفة التخصصية في المجال الموضوعي ، والنمط أو الأسلوب المعرفي ، وكيفية تأثير هذه الخصائص على مختلف مقاييس المخرجات. وتقدم بورجمان (1986) Borgman مراجعة علمية مستفيضة للدراسات المتصلة بهذا الموضوع .

وقد تعرض العديد من الدارسين لأثر الخبرة في إجراء عمليات البحث على فعالية الاسترجاع ، مقيسة بناء على الاستدعاء والتحقيق والوقت المستنفد في إجراء البحث . وفي تجربة مبكرة في هذا الموضوع ، عرّفت فينيشل (1981) Fenichell الخبرة بأنها :

إما (١) الخبرة بآليات البحث على الخط المباشر بوجه عام ، وإما (٢) الخبرة في عرض بيانات إرك ERIC الذي كان بؤرة الاهتمام في دراستها . وقد تبين من تحليل سجلات عمليات البحث أن الفروق بين المجموعات التي تعرضت للدراسة لم تكن كبيرة. كما تبين لفينيشل أن أداء مسئول البحث المبتدئين كان مرتفعاً على عكس ما كان متوقعاً ، بالمقارنة بمسئولي البحث المتمرسين ، وأن أكثر مسئول البحث خبرة في كل من الإجراءات الخاصة بنظام الاسترجاع وعرض بيانات إرك ، قد أحرزوا أعلى مستويات الاستدعاء والتحقيق ، إلا أنهم كانوا أيضاً الأوفى حظاً في مستوى ما سمي بـ «جهد البحث» (كعدد الأوامر والواصفات ، وزمن الاتصال مثلاً) مما يؤكد ، ولا غرابة في ذلك ، وجود علاقة بين جهد البحث والاستدعاء .

وقد أجريت دراسات مماثلة لدراسة فينيشل ، بواسطة كل من هوارد Howard (1982) وجيرالدين ووكر (1988) Geraldine Walker حيث قام هوارد بدراسة كل من الخبرة العامة في إجراء عمليات البحث والخبرة في التعامل مع مرصد إرك ، وتبين له أن مسئول البحث المتمرسين في إرك يجرون أفضل عمليات البحث من حيث فعالية التكلفة ويحرزون أعلى نسب التحقيق . أما جيرالدين ووكر فقد تبين لها أن المستفيدين النهائيين غير المدربين يجرون عمليات بحث بكفاءة عالية على غير المتوقع ، على الرغم من أنهم كانوا ، في دراستها ، يستنفدون ضعف الوقت المستنفد من قبل الوسطاء المتمرسين ، في إجراء عمليات البحث . ولم تجد النتائج التي انتهت إليها ووكر أي سند فيما انتهت إليه دراسة حديثة أجراها لانكستر وآخرون (1992) Lancaster et al. حيث

تبين لهم أن أعضاء هيئة التدريس وطلبة الدراسات العليا الذين يقومون بإجراء عمليات البحث في مرصد بيانات إرك المسجل على أسطوانات صوتية مكتنزة ، لم يكن بإمكانهم العثور إلا على حوالي ثلث الوثائق ذات الأهمية المحتملة بالنسبة لهم.

وفي دراسة لتأثير مختلف خصائص مسئولي البحث الأخرى قامت بلأردو Bellardo (1985) بإجراء اختبار مسبق لمبجوثيها في براعة البحث على الخط المباشر (بأدائهم في عمليتي بحث)، ومستوى الإبداع (اعتمادا على استبيانين للتقرير الذاتي) ومستوى الذكاء (بناء على أدائهم في سجل الاختبارات اللفظية والكمية للتخرج) والسمات الشخصية (اعتمادا على استبيان للاستعداد للتعامل مع الآخرين) . ويدل ما انتهت إليه من نتائج على أنه لا يمكن رد فروق الأداء في إجراء عمليات البحث إلى الاستعدادات اللفظية والكمية ، والإبداع الفني ، والميل إلى التفكير النقدي والتحليلي ، إلا في أضيق الحدود.

ومن كفاءة مسئولي البحث ، بناء على مختلف مقاييس المخرجات ، كالأستدعاء والتحقيق ، من المنطقي أن يبدأ تقييم أداء هؤلاء بمختلف متغيرات العملية. وتنطوي المحاولات الأولى في هذا المجال على إحصاء تواتر مختلف عناصر ومكونات الأوامر ، كالعوامل البولينية أو الأوامر الخاصة بعرض التسجيلات مثلا ، وتسمى فينيشل (1981) Fenichell هذه العناصر متغيرات «جهد البحث». وعلى الرغم من أهمية ملاحظة استعمال سلاسل أو عناصر أوامر بعينها من جانب مسئولي البحث ، فإنه ربما كان الأهم من ذلك هو النظر إلى هذه الأوامر في إطار نموذج لكيفية إجراء عملية البحث. ويرجع الفضل في اقتراح هذه الفكرة إلى بيتس (1979 a) Bates الذي تتكون «خطته التنفيذية» من ١٧ خطوة لإجراء عملية البحث ، ويحاول الدارسون منذ ذلك الوقت التحقق من صحتها عن طريق الملاحظة الميدانية المباشرة (1991) Wildemuth et al., (1990) Hsieh Yee,

وهناك اتجاه آخر للبحث في متغيرات عملية الاسترجاع تتبناه بورجمان Borgman (1984) . فهي تفترض أنه يمكن لأداء المستفيدين المدربين بناء على نموذج نظري لكيفية عمل نظام الاسترجاع ، يمكن أن يكون أفضل من أداء هؤلاء المدربين ببساطة على استعمال سلاسل محددة من الأوامر. ثم قامت بعد التدريب بتحليل عمليات البحث

الاختبارية ، سعيًا وراء الخصائص أو السمات التي يمكن أن تدل على « النموذج الإيجابي » (كاستعمال المنطق البوليني أو تصفح الكشاف مثلاً) أو « النموذج السلبي » (كاستعمال العبارات بدلا من الكلمات مثلاً) وذلك في نظام الاسترجاع. وقد تبين لها بوجه عام أن أداء من تدربوا على النموذج النظري كان أفضل من غيرهم في مهام الاسترجاع المعقدة، إلا أن عددا غير متوقع من مسئولى البحث ، أيا كانت الطريقة التي درّبوا بها ، كانوا يتبعون طريقة التجربة والخطأ «العشوائية» .

الفصل الحادي عشر

الطرق الآلية لاسترجاع المعلومات

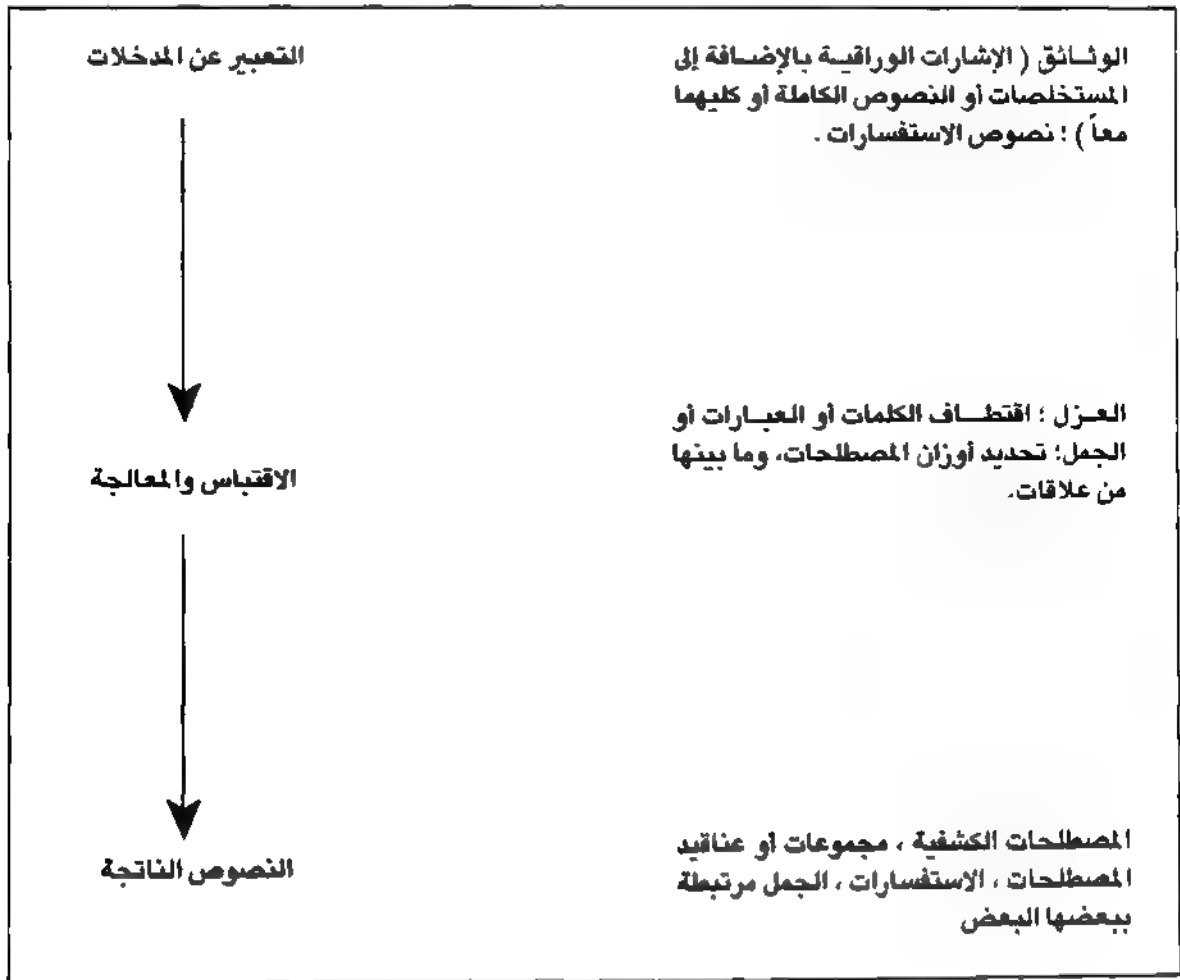
في النظم الإلكترونية التقليدية ، يتم جل إن لم يكن كل عمليات التجهيز الفكري بواسطة البشر ، بينما يكون الحاسب الآلي مجرد آلة مضاهاة عملاقة. إلا أنه من الممكن تطوير نظم استرجاع تتم فيها جميع العمليات آلياً ، حيث يتضاءل التجهيز الفكري البشري إلى أدنى حد ، وربما يتم الاستغناء عنه تماماً . فمن الممكن استخدام الحاسبات الآلية في تكشيف الوثائق ، وإعداد المستخلصات أو المقتطفات أو الاقتباسات على الأقل ، وفي تدبر استراتيجيات البحث آلياً ، ووضع الروابط بين المصطلحات المرتبطة ببعضها البعض دلاليًا ، وبذلك ينشأ شكل من أشكال أدوات البحث (نوع من المكانز يتم إعداده آلياً) . وفي النظم التي تتم فيها جميع العمليات آلياً ، يتم التكشيف بواسطة الحاسب الآلي ، حيث يتم إعداد مكنز يتكون داخلياً ، كما تتم صياغة استراتيجيات البحث آلياً ، اعتماداً على صيغة الإعراب عن الحاجة إلى المعلومات باللغة الطبيعية .

وقد بدأت تجارب الطرق الآلية في خمسينيات القرن العشرين ولا زالت مستمرة حتى الآن . وقد استخدم الباحثون الأساليب الإحصائية واللغوية ، فضلاً عن أساليب الذكاء الاصطناعي في أتمتة الإجراءات ، وفي اختبار مالها من آثار في كفاءة نظم الاسترجاع ، بناء على مقياسي الاستدعاء والتحقيق . ويلقي هذا الفصل الضوء على نماذج من الدراسات التي تصف الإجراءات التي تتبع في كل من النظم كاملة الأتمتة والنظم

التي تعتمد على الآلات ، والتي تقوم ببناء المكانز آليًا ، وكذلك الإجراءات التي تتبع في الكشف الآلي والتكشيف الذي تشارك فيه الآلات ، والصياغة الآلية للاستفسارات ، فضلا عن الاستخلاص الآلي . كذلك يستكشف هذا الفصل طرق تقييم هذه النظم ومدى ما حققته من كفاءة في الأداء.

الاتجاهات العامة للطرق الآلية لاسترجاع المعلومات :

قبل الدخول في وصف الأساليب المتبعة في أتمتة عمليات الاسترجاع نلقي نظرة عامة على المفاهيم والطرق التي استخدمت . وجدير بالملاحظة إن هذه الأساليب لا يكاد



الشكل رقم (٦١) : تصور عام للطرق الآلية لاسترجاع المعلومات.

استخدامها يتجاوز الحدود التجريبية في مجموعات صغيرة من الوثائق . ولهذا دلالاته المهمة بالنسبة لتقييم النتائج الذي نتناوله في القسم الأخير من هذا الفصل تحت عنوان «تقييم الطرق الآلية».

وكما يتبين لنا في الشكل رقم (٦١) فإنه يتم إدخال قطاع نصي (إما مجموعة من الإشارات الوراقية ، أو مجموعة من النصوص الكاملة) حيث يتم اقتطاف بعض العناصر النصية ومعالجتها بعملية آلية ، لنحصل على ناتج مركز (قوائم بالمصطلحات الكشفية أو سطور من نص مترابط). وتتوقف المخرجات على طبيعة المدخلات والإمكانات الحقيقية لعمليات التجهيز الآلي ، مما يؤكد تفاوت المستويات التي يمكن تقبلها للتجهيز. فمعظم الدراسات تستخدم الآن خوارزميات إلكترونية لاقتطاف ومعالجة النصوص الواردة في بدائل الوثائق لا أكثر ، كما حاولت أعداد قليلة من الدراسات استخدام عمليات التحليل الآلي بشكل أكثر كثافة ، وقد أصبحت هذه الفئة من الدراسات هي الغالبة في السنوات الأخيرة.

لننظر ، على سبيل المثال في الوثائق والاستفسارات الواردة في الشكل رقم (٦٢)؛ فالاستفسار الأول يتعلق باستخدام البيانات القابلة للتجهيز بواسطة الآلات (كتسجيلات مارك مثلا) كمصدر في فهرسة الوثائق. أما الاستفسار الثاني فيتعلق بالمقالات التي تتناول الأمور المتصلة بفهرسة ملفات البيانات القابلة للقراءة بواسطة الآلات . وهذان استفساران كل منهما قائم بذاته. ومن ثم فإن المستفيد الذي يهتم بالوثائق المتصلة بالاستفسار الأول قد لا يكون راغبا في الوثائق المتصلة بالاستفسار الثاني ، كما أن المهتم بالوثائق المتصلة بالاستفسار الثاني قد لا يكون راغبا في الوثائق المتصلة بالاستفسار الأول . وإذا نظرنا في الوثائق نجد أن الأولى (أ) لا تتصل بأي من الاستفسارين ، أما الوثيقة الثانية (ب) فتتناول فهرسة ملفات البيانات القابلة للقراءة بواسطة الآلات ، ومن ثم فإنها تتصل بالاستفسار الثاني. أما بقية الوثائق (جـ، د، هـ، و) فتدور حول استخدام البيانات القابلة للتداول بواسطة الآلات كمصدر للفهرسة ، وتتصل بالاستفسار الأول.

REQUESTS

"Machine-readable cataloging data"

"Cataloging machine-readable data"

PROCESSING COMPONENTS

الاستفسارات

الاستفسار الأول :

الاستفسار الثاني :

عمليات التجهيز :

cataloging
machine
readable
data

١ - اقتطاف الكلمات المفتاحية من الاستفسارين

(Using these four words, none of the records [a - f] would be retrieved because the matching algorithm does not take into account):

cataloging
cataloguing

٢ - مراجعة هجاء الكلمات .

(By substituting both spelling variants, all records [a - f] would be retrieved; however, this does not distinguish between document b, relevant to the second request, and document d, relevant to the first. To make this distinction, additional linguistic information is needed):

machine-readable cataloging data
cataloging machine-readable data

٣ - ربط الكلمات ببعضها البعض .

(By constructing different phrases, the distinction between these two expressions is maintained; however, it is still necessary to account for other, synonymous terms that might occur, as with document e. To do this, the meanings of expressions must be accounted for):

machine-readable cataloging
synonym - MARC

٤ - المعلومات الدلالية .

(By including synonymous expressions, documents containing completely different words that mean the same thing can be retrieved; however, there are other, contextual clues, which can also point to relevant documents):

٥ - المعلومات السياقية .

... libraries seeking a machine-readable cataloging data source. They include on-line systems, CD-ROM systems...

(" They," in document d, refers to " machine-readable cataloging sources, " which are the topic of the discussion in document f).

وثائق (من مرصد بيانات إرك)

١

Title: Linking approval plans and automated library acquisitions systems**Author (S):** Smith Scott A.**Journal:** Library Acquisitions: Practice and Theory**Source:** 11 (3) 1987, 215-216**Languages:** English

Abstract: Contribution to a special section devoted to the Midwinter Meeting of the American Library Association (ALA), Chicago, Jan 87. Paper presented to the RTSD RS Automated Acquisitions/ In-Process Control Systems Discussion Group. Describes the process whereby an approval vendor provides machine-readable records corresponding to books sent to a library automatically and on approval, for loading into the library's automated acquisitions system. At present vendors are offering 2 alternatives, either an LCMARC cataloguing record, or an abbreviated MARC format bibliographic record created by the vendor. Often LC records are CIP (Cataloguing in Publication Data) level, since that is usually all that is available when the vendor ships a book. The supply of records as a feature of an approval programme introduces a new aspect of customer service.

ب

Title: Cataloguing machine-readable data files: an introduction**Author (S):** Chang, Roy**Journal:** Journal of Educational Media Science**Source:** 19 (1) Autumn 81, 19-30. illus. 11 refs**Languages:** English

Abstract: Due to lack of bibliographic control and unfamiliarity, machine-readable data files (MRDFs) remain phantom material to both librarians and library users--even though more and more research materials are being produced in machine-readable form. Reviews the current status of MRDFs; discusses basic conditions for cataloguing and reviews Chapter 9 of the Anglo-American Cataloguing Rules (2 nd edition) which is devoted to descriptive cataloguing rules for MRDFs, focusing on troublesome areas.

ج

Title: Is CD-ROM a threat to bibliographic networks?**Author (S):** Fokker, Dirk W.**Editors:** Edited by Ching-chih Chen and David I. Raitt, MicroUse Information and FID, 1989**Journal:** In: Proceedings of the 2nd Pacific Conference on New Information Technology for Library & Information Professionals, Educational Media Specialists & Technologists. Singapore, 29-31 May 1989 **Source:** 93-101. 17 refs.**Languages:** English

Abstract: Bibliographic networks concentrate on cooperative and shared cataloguing from a large central data base. The provision of machine-readable cataloguing records and other cataloguing products is one of their main functions. However, libraries may experience the participation in networks as a loss of autonomy. CD-ROM offers an alternative today. With data bases available on CD-ROM containing cataloguing information for library items, libraries can buy or subscribe to these discs which will give them a high retrieval rate. CD-ROM fits easily into traditional library practices and does away with many of the problems experienced with networks. The records can be manipulated, adapted and downloaded into local systems. The consequences of using CD-ROM as a source of cataloguing data are that the basis for cooperation can affect the viability of bibliographic networks.

د

Title: CD-ROM as a cataloguing tool

Author (S): Saunders, Laverna M.

Journal: Technical Services Quarterly

Source: 5 (7) 1985, 45-58, illus. bibliog.

Languages: English

Abstract: A CD-ROM bibliographic system can be a valuable asset to larger libraries planning retro-spective conversion and to smaller libraries seeking a machine readable cataloguing data source. They include on-line systems, CD-ROM systems, or combinations of the two. The selection of a CD-ROM system should be integrated into a library's overall automation plan. Numerous questions regarding the vendor, the data base and the system's features need to be asked during the investigation process. A case study for approaching these issue is the experience of Nevada academic and public libraries with General Research Corporation's Laser Quest.

هـ

Title: Thoughts on the MARC system at BLHSS

Author (S): Bruce, Alastair

Journal: MARC Users' Group Newsletter

Source: 11 (2) Aug 87, 4-6

Languages: English

Abstract: A fairly recent recruit ■ the British Library Humanities and Social Sciences cataloguing section describes his expriences of using MARC cataloguing records in an environment of ambiguity and change. N.L.M.

و

Title: Comparison of three CD-ROM cataloguing tools: BibiloFile, LaserCat, LaserQuest

Author (S): Giesbrecht, Walter

Journal: School Libraries in Canada

Source: 9 (1) Fall 88, 23-27. table. 5 refs

Languages: English

Abstract: BiblioFile, LaserCat and LaserQuest are sources of catalogue records, they are not catalogues in themselves. All 3 Products contain MARC records in the US MARC format, the number and type of record varying however. Illustrates and compares the features of the 3 products covering searching, displaying and printing of records, original cataloguing, costs and hardware requirements. The fact that LaserCat contains the holdings of many small public libraries and some regional school libraries makes it a good source of records for many of the items which school libraries are likely to have. BiblioFile and LaserQuest, on the other hand, would seem to be best suited to libraries undergoing large retrospective conversion projects. P.B.

وفي أبسط الطرق الآلية ، يمكن للمستفيد أن يُدخل سلسلة من الكلمات المفتاحية، التي يمكن أن تستخدم للبحث في الكشافات المصنفة التي تضم الكلمات المفتاحية الواردة في الوثائق ، واسترجاع الوثائق التي تضاهاي جميع الكلمات أو العناصر التي أدخلها المستفيد . ونجد هذا الأسلوب في العديد من واجهات التعامل على الخط المباشر مع كشافات مقالات الدوريات وفهارس المكتبات المتاحة على الخط المباشر على السواء. إلا أن كل ما أمكن تقديمه ألياً هو أداة العطف المنطقية «و» التي تربط بين الكلمات المفتاحية ، وفيما عدا ذلك فإن الآلة ما زالت تعمل كأداة للمضاهاة لا أكثر. وغالباً ما تسفر هذه الطريقة الآلية البسيطة عن أخطاء خطيرة في الاسترجاع. وهذه في واقع الأمر ليست أخطاء في الخوارزمية ، حيث تقوم الآلة بتنفيذ الدور المحدد لها بدقة ، وإنما تحدث هذه الأخطاء نتيجة لعاملين :

(١) أن الخوارزمية تعتمد فقط على وجود أو عدم وجود الكلمات المحددة باعتبارها التعبير المنطقي عن الطلب الذي قدمه المستفيد لإجراء البحث .

(٢) أن تمثيل الوثيقة هنا يتكون من جميع الكلمات التي تعبر عن محتواها ، حيث لم تتم المفاضلة بين هذه الكلمات لاختيار تلك الكلمات التي يمكن أن تكون دالة أكثر من غيرها على المحتوى الموضوعي للوثيقة. وقد أدخلت بعض الطرق الآلية التي تحاول حل مثل هذه الأنواع من المشكلات . ولتحقيق ذلك فإن مثل هذه الطرق:

(١) تستخدم بعض الأساليب التي تعدل أو تدعم الكلمات المفتاحية .

(٢) تتحقق من العناصر التي تدل على محتوى الوثيقة أو الاستفسار ، وتختار هذه العناصر.

وتأخذ هذه الطرق في حساباتها الخصائص الإحصائية واللغوية للنص.

الأساليب اللغوية لاسترجاع المعلومات :

توضح الأمثلة الواردة في الشكل رقم (٦٢) بعض الأساليب والقضايا المرتبطة بالطرق اللغوية للتحليل الآلي. ويشتمل الشكل على استفسارين وعينات من الوثائق المتصلة بهما ، ويقدم أمثلة لأربعة عناصر أو مكونات وهي العنصر الصرفي والعنصر النظمي والعنصر الدلالي ، والعنصر التداولي Pragmatic ، ويمكن لهذه العناصر أن تكون جزءاً من التحليل اللغوي الآلي في استرجاع المعلومات . ويستخدم الشكل رقم (٦٣)

هذه المستويات الأربعة (بالإضافة إلى مستوى خامس وهو المستوى المعجمي Lexical) لمقارنة المكونات الأساسية في معظم نظم الاسترجاع المتاحة تجارياً ، بما يقابلها من مكونات في الأساليب اللغوية وتلك المعتمدة على المعرفة ، التي مازال معظمها في المرحلة التجريبية . ويركز هذا القسم على المكونات المألوفة في النظم المتاحة تجارياً مقارنة بتلك النظم التي تعتمد على التجهيز اللغوي الحقيقي أو تعتمد على المعرفة . أما الطرق الإحصائية فنتناولها في القسم التالي.

فمن الكلمات المفتاحية الواردة في استفساري البحث ، في الشكل رقم (٦٢) كلمة "Cataloging" إلا أن مرصد البيانات كما ورد في الشكل التوضيحي يستعمل شكلاً هجائياً آخر "Cataloguing" . وهذه من المشكلات التي يحلها المستفيدون بإدخال الأشكال الهجائية المختلفة أو باتباع مبدأ البتر . وغالباً ما يكون البرنامج في النظام الآلي قادراً على البتر وتقديم الأشكال الهجائية المختلفة.

وفي التحليل الصرفي الحقيقي ، يتم تحليل الجذوع والصدور والكواسع لتحديد نوع الكلمة (ما إذا كانت اسماً أم فعلاً أم نعتاً ، وهكذا) . ويتم اختزان المعلومات حول كل كلمة على حدة ونوعها في معجم Lexicon يتم الرجوع إليه لتجهيز كل عنصر تجهيزاً صحيحاً ، وتستخدم هذه البيانات كمدخلات لنظام فرعي للإعراب Parser ، وهو برنامج حاسبي يقوم جزئياً أو كلياً بتحليل البنية اللغوية للجملة . وعلى ذلك ، فإنه في هذه الحالة يتم الاحتفاظ بصدور الكلمات وكواسعها لتقديم المعلومات اللازمة للتحليل اللغوي . فمن الممكن على سبيل المثال للنظام الفرعي للتحليل أن يسم في هذه المرحلة الكلمات التالية كما هو مبين :

| | | |
|------------------|-------|--------------|
| Comput - e | V | (فعل) |
| Comput - er | N | (اسم) |
| Comput - ed | PPart | (ماضٍ تام) |
| Comput - ational | Adj | (نعت) |

ويتطلب الأمر هذا النوع من التحليل إذا كان من المزمع تحليل نص المدخلات إلى وحدات مركبة كالعبارات الاسمية المركبة ، كما سيرد وصفه فيما بعد.

وبمجرد حل مشكلة اختلاف الشكل الهجائي أو الصرفي للكلمات ، يتعين النظر في المشكلات الأخرى ، فاستفسارا البحث عندما اختزلا في شكل كلمات مفتاحية أصبحا متطابقين ، كما أن هذين الاستفسارين ، كانا في البداية ، حين كانا عبارات فعلا ، متميزين في المعنى ، إلا أنهما أصبحا الآن غامضين نتيجة لاختزالهما في شكل قائمة موحدة من الكلمات المفتاحية . ومن ثم فقد تم استرجاع الوثيقتين « ب » و « د » لكل من الاستفسارين ، بينما تتصل كل واحدة منهما باستفسار واحد فقط . ويمثل هذا القصور في دقة الاستفسار الآن إحدى المشكلات نظراً لأن الاختلاف في بنية العبارات قد تلاشى . ويعمل المستفيدون على حل هذه المشكلة بربط الكلمات في عبارات كاملة اعتماداً على تعليمات البحث اللازمة . وفي النظام الآلي يتم تحديد العبارات الكاملة في النص ومعالجتها اعتماداً على شكل من أشكال التحليل النظمي .

هذا ، ولم يتطلب حل مشكلة العبارات أكثر من مجرد إدخال بعض التعديلات على الكلمات المفتاحية الأصلية ، ومن ثم فإن الاسترجاع ما زال عملية مضاهاة مع بعض التطوير الإضافي في شكل البتر الآلي (رد الكلمات إلى جذورها) فضلاً عن تحليل العبارات . إلا أن الوثيقة « هـ » ما زال من غير الممكن استرجاعها استجابة للاستفسار الأول ، ولما كانت هذه الوثيقة تبدو صالحة ، فإن ذلك يعني وجود قصور في الاستدعاء . ولكي يتم استرجاع الوثيقة « هـ » فإنه يتعين على المستفيد تقديم مصطلحات أخرى مرادفة أو شبه مرادفة ، يمكن أن تستعمل في عملية المضاهاة . وهذه المصطلحات ليست أشكالاً صرفية أو نظامية أخرى من الكلمات أو العبارات نفسها ، وإنما كلمات وعبارات جديدة ، قصد بها أن تكون مطابقة أو مشابهة للمصطلح الأصلي في المعنى . ولتنفيذ ذلك ألياً فإنه ينبغي أن يتوافر للحاسب الآلي إمكانية التعامل مع مصدر يتم فيه اختزان واسترجاع المعلومات المتصلة بالمصطلحات وما بينها من علاقات دلالية بشكل كفاء .

وأخيراً ، فإنه على الرغم من توافر بعض المعلومات الدلالية ، من الممكن أن نخطيء بعض الوثائق الصالحة . وإذا أدرك أحد المستفيدين ذلك ، فإنه غالباً ما يستعمل

المعلومات السياقية المستمدة من إحدى الوثائق الصالحة لكي يتوسع في البحث عن المزيد من الوثائق المشابهة. أما في النظام الآلي فإنه يتم الاستعانة ببعض عناصر الاتصال الإضافية الواردة في المستخلص للإشارة إلى وثائق صالحة أخرى. ففي حالة الوثيقة « د » ، فإن الضمير « They » يعود على « بيانات الفهرسة القابلة للقراءة بواسطة الآلات » متبوعة بأنواع بعينها ، من بينها نظم الأسطوانات الضوئية المكتنزة . ويمهد الاستدلال أو الاستنتاج بأن نظم الأسطوانات الضوئية المكتنزة من « بيانات الفهرسة القابلة للقراءة بواسطة الآلات » الطريق للوصول إلى الوثيقة « و » الصالحة أيضاً بالنسبة للاستفسار الأول.

وفي علم اللغة ، فإن الأنواع الخمسة من المعلومات التي يمكن تحديد معالمها ومعالجتها في نظام آلي ، تنقسم إلى فئات أو مستويات تقابل مستويات التعقد المتزايد بالنسبة للنظام الآلي . وهذه المستويات هي المعجمي والصرفي (بنية الكلمة) والنظمي (بنية العبارة والجملة) والدلالي (المعنى) والتداولي (اللغة في السياق) . ومن الملاحظ وجود توازن بين ما يتم إجراؤه آلياً وما يقدمه مسئولو البحث عادة يدوياً عند استخدام النظم المتاحة تجارياً. وقد حظي هذا التناظر بالمناقشة والتوضيح من قبل كل من دورزكوكس (1986) Doezkocs و وورنر (1992) Warner ، كما يكشف عنه الشكل رقم (٦٣).

| المستوى اللغوي | استرجاع المعلومات المتاح تجارياً | استرجاع المعلومات الإحصائي | استرجاع المعلومات اللغوي / القائم على المعرفة |
|--|---|---|--|
| المعجمي الصرفي النظمي الدلالي | قائمة الكلمات المستبعدة رمز البتر معاملات التقارب المكنز | قائمة الكلمات المستبعدة التجريد العبارات الإحصائية عناقيد الكلمات التي ترد معاً | المعجم التحليل الصرفي العبارات النحوية شبكة الكلمات / العبارات المرتبطة ببعضها البعض دلاليًا |
| السياقي | الحواجز أو الحدود | التلقيح المرتد من الصلاحية | الواجهات الخبيرة ذات الأساليب الإيعازية في البحث |

الشكل رقم (٦٣) : المقارنة بين مستويات التجهيز في الطرق التجارية والإحصائية واللغوية لاسترجاع المعلومات .

لقد اقتصرتم الأمثلة السابقة على لغة طلبات البحث والوثائق ، بينما يعتمد النظام الآلي على برنامج حاسبي يقوم بمعالجة الاستفسارات والوثائق على مستويات

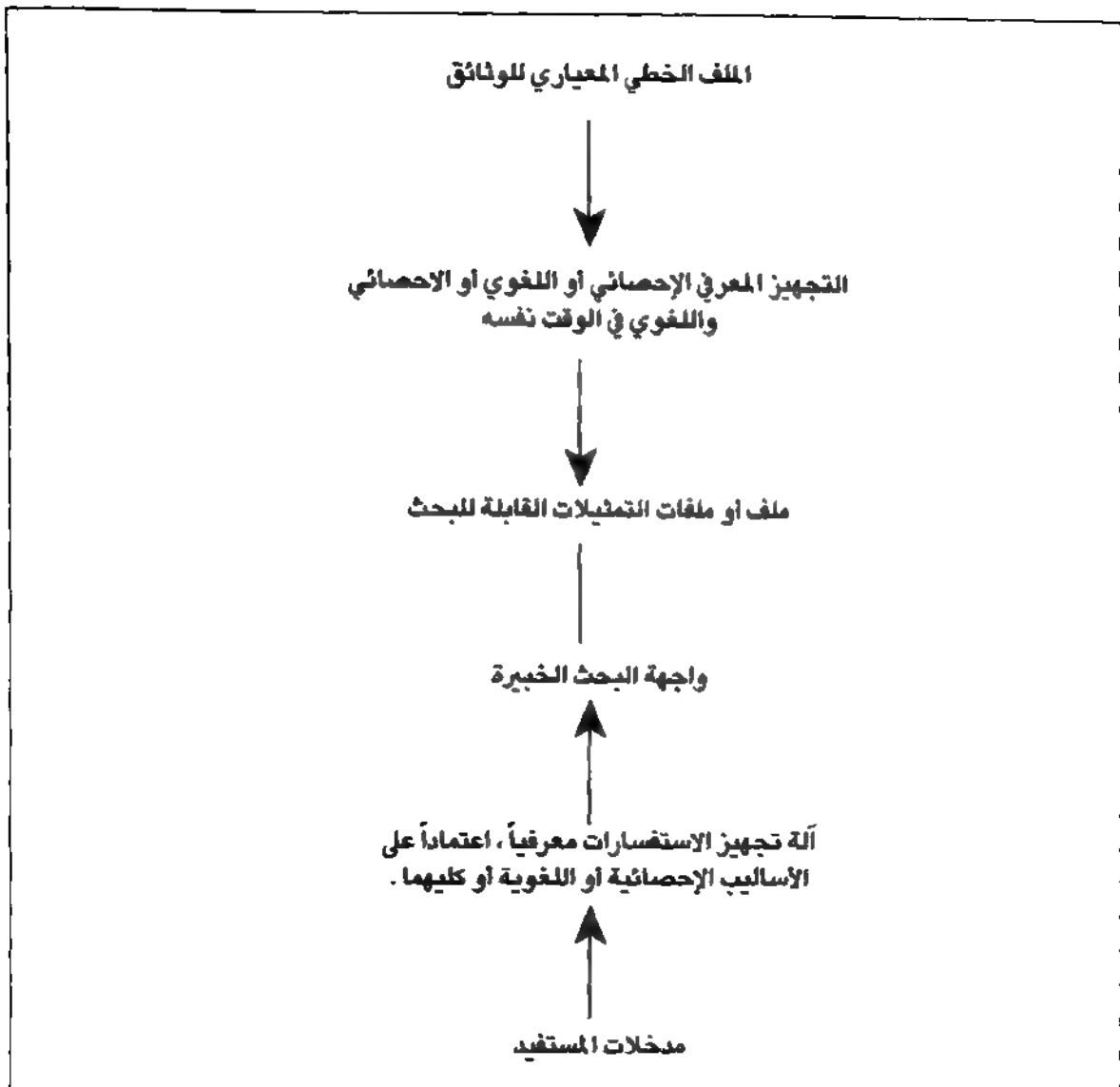
مختلفة؛ وهناك المعرفة الإضافية التي يأتي بها الفرد إلى عملية البحث ، والتي تشمل أنواعاً متعددة من إيعازيات البحث ، كالقواعد المألوفة التي تستخدم لتعديل البحث بمختلف الطرق (Fidel, 1984,1991) . وتشمل إيعازيات البحث أدوات الاستدعاء (كاستعمال المصطلحات الإضافية المرتبطة ببعضها البعض بعلاقة «أو» واستعمال الكلمات المفتاحية المفردة ، على سبيل المثال لا الحصر) ، وأدوات التحقيق (كالتجميعات المؤتلفة Combinations من المصطلحات المرتبطة ببعضها البعض بعلاقة «و» ، واستعمال العبارات الكاملة ، على سبيل المثال) . ومن ثم فإن الأمر يتطلب في الأساس استخدام نوعين من المعرفة :

(١) التحقق من البنى اللغوية العادية .

(٢) استخدام هذه البنى لتحقيق غايات استرجاعية معينة .

ويوضح الشكل رقم (٦٤) الهيكل العام لمثل هذا النظام ، كما يبين القطاعات أو النظم الفرعية modules المختلفة اللازمة لتحليل لغة الوثائق والاستفسارات ، كما سبق أن عرضنا لها ، وكذلك المعرفة الخبيرة الإضافية الخاصة بعملية البحث والتي يمكن لمثل هذا النظام أن يشتمل عليها . ويمكن لنظام من هذا النوع أن يكون نظاماً خبيراً حقيقياً للبحث ، يجعل منه أحد تطبيقات الذكاء الاصطناعي . وقد حظيت محاولات بناء مثل هذه النظم بالوصف في الإنتاج الفكري لاسترجاع المعلومات ، من جانب كل من بولت (1987) Pollitt وفيكري (1988) Vickery .

وإذا ما أمكن عزل كل هذه المعرفة واستعمالها بمستوى التطور والمرونة والنجاح نفسه الذي يعمل به مسئول البحث المتمرس ، فإنه يمكن حينئذ للنظم الآلية أن تقضي على الحاجة إلى الوسيط البشري إلا أن التطبيقات اللغوية والنظم الخبيرة مازالت باهظة التكلفة حاسبياً ، حيث تتطلب كتابة برامج معقدة خاصة بالمستويات اللغوية الخمسة ، وكذلك من أجل استخدام إيعازيات البحث . هذا بالإضافة إلى أن معظم النظم المتطورة ينبغي أن تشتمل على قواعد معرفية معقدة للمعلومات الدلالية والتداولية أو السياقية ، والتي يتسم بناؤها بالصعوبة فضلاً عما يتطلبه من وقت طويل . ومما يدعو للعجب في النهاية أن هذه الطرق لم تحقق سوى مظاهر تحسن متواضعة في كفاءة النظم ، على الرغم من مغرياتها البديهية .



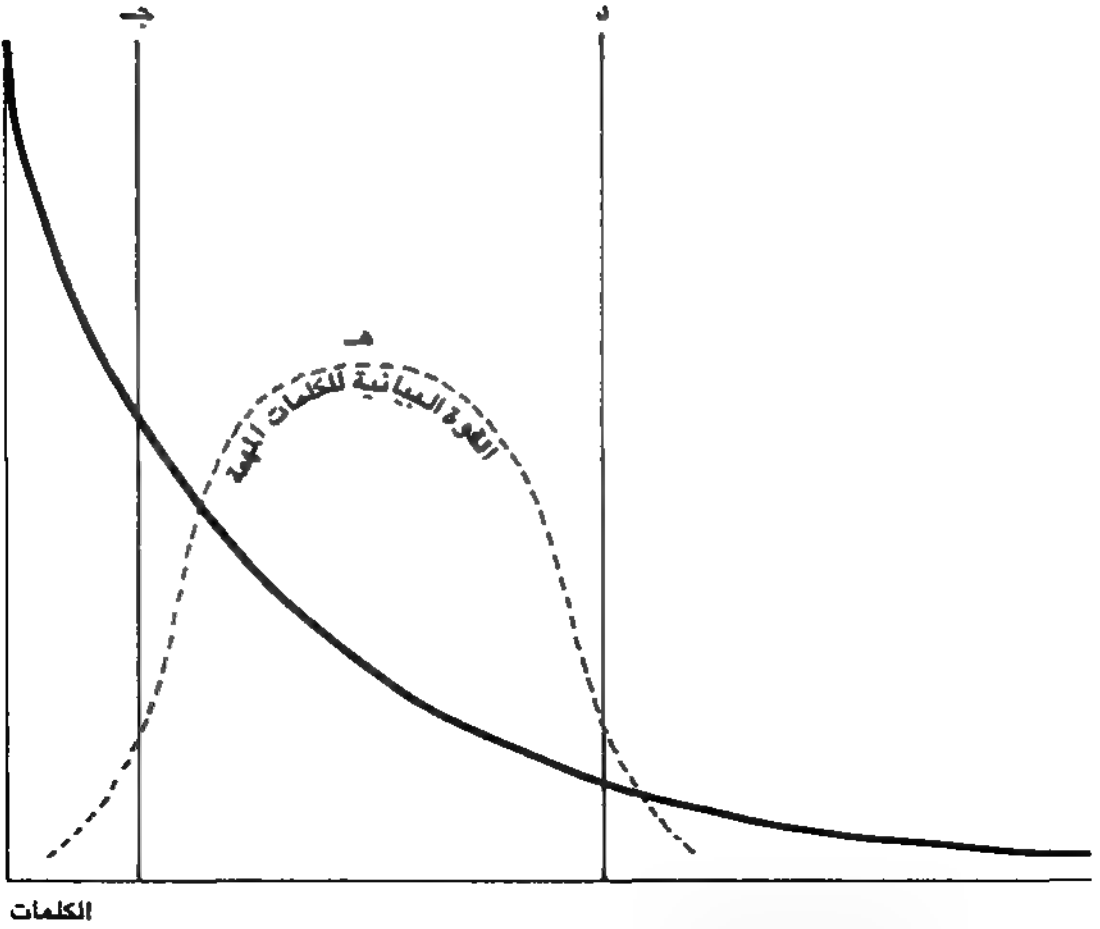
الشكل رقم (٦٤) : هيكل مبسط لنظام ذكي لاسترجاع المعلومات .

الطرق الإحصائية لاسترجاع المعلومات :

يعتمد قطاع كبير عريق من الجهود المتصلة بالطرق الآلية لاسترجاع المعلومات على بعض الخصائص الإحصائية للعناصر النصية الواردة في الوثائق (انظر العمود الثالث في الشكل رقم ٦٢). وتستند هذه الجهود إلى دراسات زيف (Zipf (1935 الذي لاحظ بين ما لاحظ أن :

١ - الكلمات القصيرة تستعمل في الإنجليزية بكثافة أكثر من الكلمات الطويلة :

« فأيما كانت ضخامة رصيد الكلمات القصيرة والكلمات الطويلة فإن الشواهد اللغوية تدل بجلاء لا لبس فيه على أنه كلما ازدادت الكلمة طولاً تضاءلت احتمالات استعمالها » . (P. 22) .



الشكل رقم (٦٥) : القطاع (ج) الأعلى والقطاع (د) الأدنى لتواتر أو تردد الكلمات.
وتتمتع الكلمات ذات التردد المتوسط بقوة بيانية ملائمة .

٢ - يميل توزيع تردد استعمال الكلمات في الإنجليزية للانحراف ؛ فهناك قليل من الكلمات عالية التردد ، وكثير من الكلمات قليلة التردد : « وكلما زاد عدد مرات التردد تناقص عدد الكلمات المختلفة التي تتردد بذلك المعدل » . (P.41)

وتفاوت توزيع تردد هياكل الكلمات تبعاً لمختلف أنواع الكلمات ظاهرة أساسية تستند إليها كثير من الطرق الإحصائية المستخدمة في الاسترجاع الآلي للمعلومات ، بدءاً بجهود هانز بيتر لون (Hans Peter Luhn 1957,1958) في نهاية خمسينيات القرن العشرين . وقد انتهى لون إلى أنه من الممكن بعد ترتيب جميع الكلمات الواردة في مجموعة ما أو في عينة من مجموعة ، طبقاً تبعاً لعدد مرات تردها ، من الممكن تحديد

معالم قطاعين ، أحدهما في الجانب الأعلى حيث تنقسم الكلمات بالآلفة أو الشيوخ إلى الحد الذي تصبح معه عديمة الجدوى في تمييز مجموعات الوثائق ، وثانيهما في الجانب الأدنى حيث تنقسم الكلمات بعدم الشيوخ مما يجعلها عاجزة عن الدلالة على مجموعات الوثائق أو فئات الوثائق بشكل يمكن الاعتماد عليه. ويسمى لون هذه الخاصية من خصائص الكلمات «بالقوة البيانية» أو «درجة التمييز» (Luhn, 1958, P. 160) . ويوضح الشكل رقم (٦٥) المقتبس من مقالة لون (Luhn, 1958, P. 161) هذا الموقف ، حيث يحد الخط «ج» قطاع التردد المرتفع والخط «د» قطاع التردد المنخفض، وحيث تعد الكلمات التي تتردد في المدى المتوسط «هـ» هي الكلمات التي تتمتع بقوة بيانية مناسبة.

ونجد أكثر تطبيقات مبادئ التردد هذه شيوعاً في كل من النظم الإحصائية الآلية والنظم التجارية غير المعتمدة على الأساليب الإحصائية، في استخدام قائمة الاستبعاد لتنحية الكلمات الوظيفية عالية التردد من تسجيلات الوثائق والاستفسارات ، واستبقاء كلمات المضمون دون غيرها لتكون محل الاعتبار. ويتم تجهيز كلمات المضمون ، في معظم النظم الإحصائية ، باستعمال برنامج يعرف «بالمجرد» وهو مناظر للمحلل الصرفي ، وإن كان أقل منه تطوراً وتعقداً. وهنا تتم تنحية جميع اللواحق كلية عن الكلمات بحيث تصبح عبارة عن «أصول» أو «جذوع» ، أي يصبح من الممكن مثلاً اختزال كل من Comput [ational] Comput [er] Comput [ed] في شكل واحد أو صيغة واحدة، ثم تحول هذه الأشكال بعد ذلك للتحليل الإحصائي.

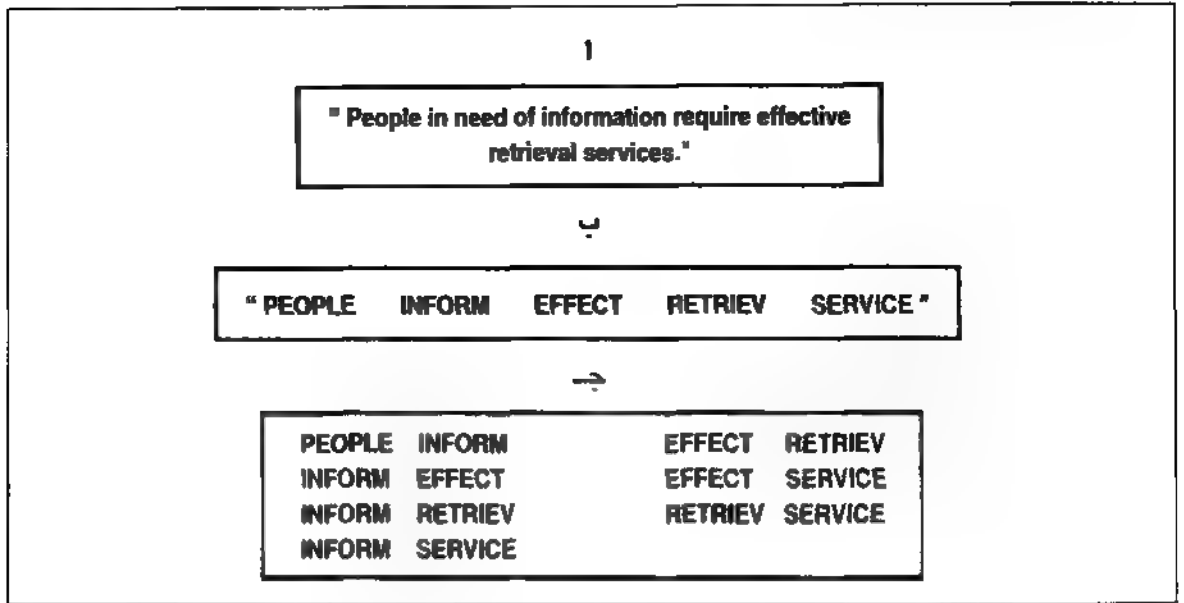
ويمكن لبرامج التجريد أن تعمل بعدة طرق يلخصها سالتون (1989) على النحو التالي:

- ١ - بإمكانها تنحية كواسع وصدور كلمات معينة.
- ٢ - بإمكانها تنحية عدد محدد من الحروف الواردة في نهاية الكلمات.
- ٣ - بإمكانها اختزال كل كلمة إلى جذع بحجم محدد.

وعادة ما تستعمل الطرق الآلية التي تنطوي على إحصاءات، الكلمات المفتاحية فقط في تحليلاتها . إلا أن هناك عدداً قليلاً من الطرق التي تتعامل بالعبارات ، وعادة ما يتم في هذه النظم تحديد العبارة إحصائياً لا لغوياً. والهدف هنا هو التقاط الكلمات ذات

الترددات المرتفعة جدا وربطها بكلمات أخرى ، بحيث تتكون عبارات ذات ترددات متوسطة .و يصف كل من سالتون وماكجل (Salton and Mc Gill (1983 هذا الإجراء في نظام سمارت SMART :

- ١ - تستخدم قائمة الاستبعاد لتنحية الكلمات الوظيفية العامة.
 - ٢ - يستخدم برنامج التجريد للحصول على جذوع الكلمات.
 - ٣ - التقاط ثنائيات من جذوع الكلمات الناتجة ، بحيث يشكل كل ثنائي عبارة ، بشرط ألا تزيد المسافة في النص بين عنصري العبارة عن حد معين ، وأن يكون أحد مكونات كل عبارة على الأقل من الكلمات عالية التردد. (P.133)
- ويتم الحصول على ثنائيات الكلمات التي تعد بمثابة عبارات في هذا النظام ، من الجمل التي يتم إدخالها ، والتي يوضح الشكل رقم (٦٦) مثالا لها. ومن الممكن مقارنة هذه الطريقة بالطريقة اللغوية لتحليل العبارات والتحقق منها ، كما هو مبين في الشكل رقم (٦٧).



الشكل رقم (٦٦) : مثال لصياغة العبارات إحصائياً : (أ) الجملة الأصلية : (ب) جذوع الكلمات الواردة في الجملة الأصلية : (جـ) ثنائيات الكلمات (العبارات) الناتجة بناء على أن الفاصل بين ثنائيات الكلمات لا يزيد على ثلاث كلمات .

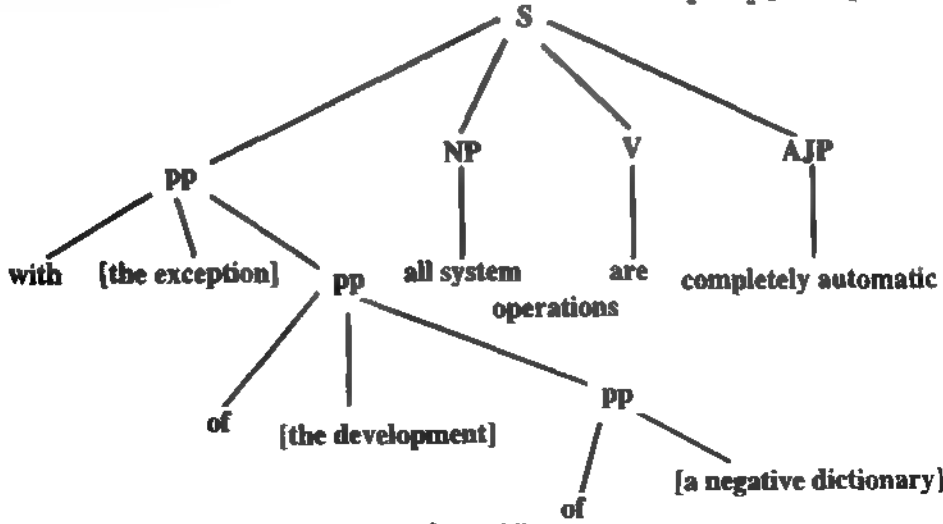
ويتم انتقاء أو تنظيم العبارات المفردة أو الجذرية الناتجة عن هذه العملية تبعاً لسلوكها الإحصائي ، بطريقتين مألوفتين في الاستخدام : (١) «التردد المطلق» حيث يتم اختيار أو تجميع الجذوع أو العبارات المفتاحية التي ترد « س » من المرات أو أكثر.

و (٢) «التردد النسبي» حيث يتم اختيار أو تجميع الجذوع أو العبارات المفتاحية عندما تتردد في وثائق معينة بمعدل أعلى مما كان متوقفاً بناءً على ترددها في المجموعة ككل .

عينة جملة من مستخلص :

" With the exception of the development of a negative dictionary, all system operations are completely automatic."

تسفر عن التحليل النحوي التالي :



الذي يسفر عن العبارات التالية :

Development exception
Dictionary development
Negative dictionary
System operations

S = sentence
PP = prepositional phrase
NP = noun phrase
V = verb
AJP = adjectival phrase

جملة
عبارة بها حرف جر
عبارة اسمية
فعل
عبارة نعتية

الشكل رقم (٦٧) : عينة لعبارة كشفية ناتجة عن التحليل النحوي.

وفي واحد من أشهر النظم الآلية وأكثرها عراقية ، وهو نظام سمارت SMART (Salton, 1968) يتم التقاط كل كلمة من كلمات المضمون الواردة في عبارات الاستفسارات وفي الوثائق ، وتجريدها ، وتحديد أوزانها تبعاً لتردها في الاستفسارات والوثائق . ثم يتم بعد ذلك المقارنة بين مجموعات الجذوع الموزونة الناتجة من الاستفسارات والوثائق وحساب معامل التشابه ، حيث يدل « ١ » على التطابق التام و « صفر » على الاختلاف التام . ويشتمل الشكل رقم (٦٨) على مثال لهذه العملية مقتبس من كتاب سالتون.

GAVE ALGORITHMS USEFUL FOR THE NUMERICAL SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS AND PARTIAL DIFFERENTIAL EQUATIONS ON DIGITAL COMPUTERS. EVALUATION THE VARIOUS INTEGRATION PROCEDURES (TRY RUNGE-KUTTA, MATRIX METHOD) WITH RESPECT TO ACCURACY, STABILITY, AND SPEED.

2 2 2 2 2

[illegible]

TEXT MICROPROGRAMMING

SMICRO-PROGRAMING
DR. J. MEYER (UNIVERSITY OF CALIFORNIA)
SUN. GOV. RES. REPTS. VOL. 20 PP 71-72 AUGUST 15, 1984 PG 126833

MICRO-PROGRAMMING. THE MICRO-PROGRAMMING TECHNIQUE OF DESIGNING THE CONTROL CIRCUITS OF AN ELECTRONIC DIGITAL COMPUTER TO FORMALLY INTERPRET AND EXECUTE A GIVEN SET OF MACHINE OPERATIONS AS AN EQUIVALENT SET OF SEQUENCES OF ELEMENTARY OPERATIONS THAT CAN BE EXECUTED IN ONE PULSE TIME IS DESCRIBED.

TEXT THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS

THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS
 SM. M. ASTRAHAN (IBM CORP.)
 SIM J. REZ. AND DEV. VOL. 2 PP 310-313 (OCTOBER 1984)

THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS. THE ROLE OF LARGE MEMORIES IN SCIENTIFIC COMMUNICATIONS IS DISCUSSED. LARGE MEMORIES PROVIDE AUTOMATIC REFERENCE TO MILLIONS OF WORDS OR MACHINE-READABLE CODED INFORMATION TO MILLIONS OF IMAGES OF DOCUMENT PAGES. HIGH DENSITIES OF STORAGE WILL MAKE POSSIBLE LOW-COST MEMORIES OF BILLIONS OF WORDS WITH ACCESS TO ANY PART IN A FEW SECONDS OR COMPLETE SEARCHES IN MINUTES. THESE MEMORIES WILL SERVE AS INDEXES TO THE DELUGE OF TECHNICAL LITERATURE WHEN THE PROBLEMS OF INPUT AND OF THE AUTOMATIC GENERATION OF CLASSIFICATION INFORMATION ARE SOLVED. DOCUMENT FILES WILL MAKE THE INDEXED LITERATURE RAPIDLY AVAILABLE TO THE SEARCHER. MACHINE TRANSLATION OF LANGUAGE AND RECOGNITION OF SPOKEN INFORMATION ARE TWO OTHER AREAS WHICH WILL REQUIRE FAST, LARGE MEMORIES.

وأخيرا ، يمكن أيضاً في نظام سمارت جعل هذه العملية تكرارية ، بإجراء يعرف «بالتقييم المرتد للصلاحية» ، يصفه سالتون (1968) Salton كما يلي :

«تتاح للمستفيد فرصة الاطلاع على بعض المخرجات الأولية والتحقق من صلاحية أو عدم صلاحية بعض الوثائق بالنسبة له ، ثم يقوم النظام بعد ذلك آلياً بتعديل طلب البحث بزيادة وزن مصطلحات الطلب ، والتي وردت أيضاً في مجموعة الوثائق الصالحة كما حددها المستفيد ، في الوقت نفسه الذي يتم فيه خفض وزن مصطلحات الطلب الواردة أيضاً في الوثائق غير الصالحة» (P.17).

ولأسلوب الإحصائي للطرق الآلية مكوناته المناظرة لتلك التي عرضنا لها بالنسبة للأساليب اللغوية ، وقدراتها الموازية في النظم التجارية لاسترجاع المعلومات ، كما يوضحها الشكل رقم (٦٣).

وعلى الرغم من تقديم أمثلة بسيطة نسبياً لتوضيح الطرق اللغوية والإحصائية المستخدمة في الاسترجاع الآلي للمعلومات ، فإنه لا ينبغي التسليم بوجود أسلوب متفق عليه لاتباع أي من الطرق اللغوية أو الإحصائية. وواقع الأمر ، ونظراً لاختلاف مستويات التحليل الممكنة ، فإنه عادة ما تكون هناك بعض الطرق المستندة إلى التحليل اللغوي وبعض الطرق الأخرى المستندة إلى التحليل الإحصائي . فأحياناً ما يستخدم على سبيل المثال برنامج للإعراب (أي التحليل اللغوي) لالتقاط العبارات الاسمية النحوية من الوثائق التي تحلل بعد ذلك إحصائياً. وتنطوي النظم الآلية لاسترجاع المعلومات فعلاً على كثير من التدابير التي يمكن تطويعها بمختلف السبل ، ولكن على أمل أن يؤدي أي تعديل دوماً إلى الارتفاع بمستوى كفاءة الاسترجاع.

فئات الطرق الآلية لاسترجاع المعلومات :

هناك أربعة أنواع واضحة المعالم من الطرق الآلية للاسترجاع التجريبي للمعلومات ، وهي بناء المكانز آلياً ، والتكشيف الآلي ، وصياغة الاستفسارات آلياً ، والاستخلاص الآلي . وعلى الرغم من استقلالها عن بعضها البعض فإن هذه الأنواع جميعاً تشكل أحد مكونات نظام استرجاع المعلومات ، وكثير مما يحدث مع أي منها انعكاساته على الأنواع الأخرى . وأتمتة هذه العمليات ليست بالأمر اليسير ، نظراً لأن ما تحاوله فعلاً في هذا الصدد هو أتمتة التحليل الموضوعي ، وهي مهمة فكرية غاية في الذاتية . وربما تدفع هذه العبارة البعض للتساؤل لماذا تبذل هذه الجهود إذن؟، ويبدو أن هناك إجابات عملية وإجابات نظرية لهذا السؤال في الوقت نفسه.

فعلى الجانب العملي يعد بناء المكانز وإعداد المستخلصات والكشافات يدوياً من الأمور المبددة للوقت باهظة التكلفة . أضف إلى ذلك أنه على الرغم من أن إجراء جميع عمليات التحليل الموضوعي يدوياً ربما كان ممكناً في البدايات المبكرة لاسترجاع المعلومات ، فإنه يتوافر الآن كميات هائلة من المعلومات النصية في شكل إلكتروني ، ويتعين ترك قدر كبير منها دون أية معالجة موضوعية . واستعمال الكلمات المفتاحية في بحث النصوص الكاملة من الأمور المألوفة الآن في الاسترجاع على الخط المباشر ، كما أن بحث تلك الملفات التي تفتقر إلى المصطلحات الكشفية المعينة يدوياً أو آلياً يثير كثيراً من المشكلات في أحسن الظروف . هذا بالإضافة إلى أنه بينما كان وسطاء البحث المتمرسون يقومون غالباً بإجراء عمليات البحث على الخط المباشر نيابة عن المستفيدين في الماضي ، فإن المستفيدين النهائيين يقومون الآن بإجراء عمليات البحث الخاصة بهم ، وينبغي أن يكون البحث بالنسبة لهم عملية مباشرة واضحة لا لبس فيها . ومن ثم فإن واجهات البحث ينبغي أن تتضمن نوعاً من التجهيز المرن لما يدخله المستفيد باللغة الطبيعية ، بدلا من اللجوء دائما إلى قواعد التعامل الجامدة ، كما كان يحدث من قبل .

أما على الجانب النظري ، فإن الباحثين يحاولون استخدام الطرق الآلية للتعرف على كيفية اضطلاع نظم استرجاع المعلومات بمهامها ، فضلاً عن التعرف على ما يمكن أن يطرأ على كفاءة هذه النظم من تغير في مختلف الظروف . ولهذا الجهد أهميته لأنه يحاول أن يبين كيف يمكن للطرق الآلية أن تسفر عن تسجيلات أو بدائل للوثائق والاستفسارات لا تقل كفاءة أو أكثر كفاءة من النظم العاملة حالياً والمعتمدة على البحث البولييني للكلمات المفتاحية للغة الطبيعية والواصفات والعبارات . هذا بالإضافة إلى أن الأساليب الآلية لاسترجاع المعلومات قد شهدت اختلافات شاسعة على مر السنين ، حيث تراوحت ما بين الطرق الإحصائية البسيطة والأساليب الإحصائية المعقدة المتطورة التي تنطوي على التحليلات الأكثر تعقداً للبنى اللغوية ، فضلاً عن الإيعازيات العامة التي تنطوي عليها عمليات التكشيف والاستخلاص وبناء المكانز وصياغة الاستفسارات . وقد جاء هذا التفاوت نتيجة للخبرات والمعارف التي اكتسبها الباحثون في تعاملهم مع هذه النظم ، هذا بالإضافة إلى تطور العتاد والبرمجيات الذي أتاح إمكانية تنفيذ الخوارزميات الأكثر تعقداً من وجهة النظر الحاسوبية .

ونقارن في الأقسام التالية الطرق الآلية بنظائرها اليدوية ، وذلك في إطار سميث

ووورنر (1984) Smith and Warner الذي يتناول ما يتم تمثيله والشكل الذي يتم به التمثيل . كما نناقش أيضاً أمثلة من نتائج مختلف الإجراءات لإبراز مدى التنوع والتعدد في الأساليب.

ومن الملاحظ أن هناك فعلاً اتجاهين واضحين المعالم فيما يتعلق بما هو جدير، نظراً لأهميته ، بالتحقق في دراسات وتجارب استرجاع المعلومات. ويؤكد الاتجاهان في الأساس الاختلاف في التركيز بين الباحثين الذين يدرسون نظم استرجاع المعلومات وأولئك الذين يركزون على المستخدمين . ويركز الفريق الأول على دراسة نظم الاسترجاع في حد ذاتها ، بإدخال التعديلات على مختلف مكونات هذه النظم ثم التعرف على ما لتلك التعديلات من أثر في كفاءة الاسترجاع . ومن ثم فإنه يمكن للباحث على سبيل المثال أن يستخدم عدداً من خوارزميات الكشف الآلي ، ثم يقارن بين هذه الخوارزميات بناءً على الاستدعاء والتحقيق . أما الفريق الآخر الذي ينظر إلى الموقف من زاوية المستخدمين فيركز على كيفية تطوير واجهات التعامل التي يمكن أن تستخدم بشكل أيسر وأكثر فعالية ، وفي غالب الأحيان مع نظم الاسترجاع التي تتكون من الكلمات المفتاحية والعبارات التقليدية ، التي يتم تجهيزها من حقول النصوص الحرة ، فضلاً عن الواصفات التي يتم تعيينها . وغالباً ما تتكون واجهات التعامل هذه من مستودع ثري من الأدوات المساعدة لعمليات البحث المعقدة والتي ترمي إلى تيسير مهمة الباحث.

بناء المكانز ألياً:

يقارن الشكل رقم (٦٩) بين الأسلوبين اليدوي والآلي في بناء المكانز . وما يتم تمثيله أو التعبير عنه في كلا الحالتين مجموعة من المصطلحات الموضوعية المنظمة. وكما بينا في الفصل السادس فإن تقنين المصطلحات مفيد نظراً لأنه يحد من التنوع أو الاختلاف الناتج عن الترادف في اللغة ، هذا بالإضافة إلى أنه يدعم مقومات التنبؤ باستعمال اللغة في استرجاع المعلومات ، وذلك بالتعبير بوضوح عن كثير من العلاقات الدلالية المهمة بين الكلمات والعبارات.

ويهدف التعبير عن العلاقات الدلالية إلى تعزيز الاستدعاء في البحث بضمان كشف جميع الوثائق التي تتصل بموضوع بعينه بشكل لغوي واحد. إلا أن المكانز تسهم أيضاً في تعزيز التحقيق في البحث بمراعاة التخصيص في مصطلحاتها ؛ فكلما ازدادت

ما يتم التعبير عنه : مصطلحات موضوعية منظمة

| آلياً | | يدوياً | |
|---|--|--|------------------|
| لغوياً | إحصائياً | | |
| اختيار الكلمات الدالة / العبارات بالتحليل اللغوي (النحوي أو الصرفي) | اختيار الكلمات المفتاحية بناء على تردها (باعتبارها أدوات تمييز جيدة) | التحقق من المصطلحات بواسطة لجنة التقاط المصطلحات من الإنتاج الفكري | إجراءات الاختيار |
| استخلاص العلاقات الدلالية من النص | التجميع بناء على تردد الكلمات بصحبة بعضها البعض أو التردد النسبي | التحقق من العلاقات الدلالية | إجراءات التنظيم |
| شبكة دلالية للمصطلحات بالإضافة الى العلاقات | تجميعات عنقودية للكلمات الدالة أو العبارات الدالة | قوائم بالمصطلحات والعلاقات الدلالية | شكل التعبير |

الشكل رقم (٦٩) : مقارنة بين الطرق اليدوية والآلية لبناء المكانز.

مصطلحات المكانز تخصيصاً ازدادت احتمالات تعزيز التحقيق (إلا أن ذلك غالباً ما يكون على حساب الاستدعاء).

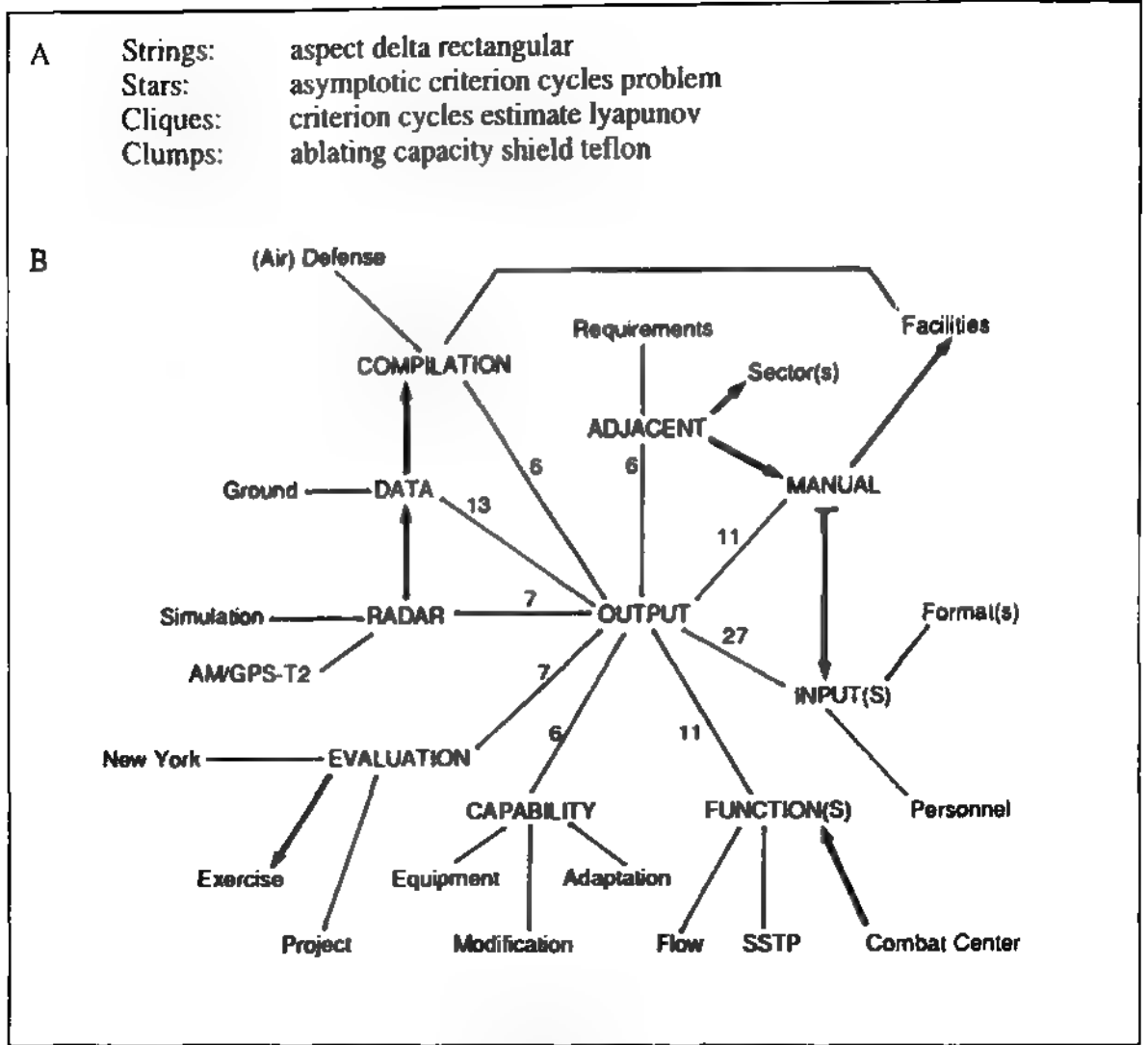
وفي الطريقة اليدوية يتم التحقق من الكلمات والعبارات بواسطة لجنة من الخبراء ، أو باستخلاصها من الإنتاج الفكري الموضوعي ، أو بكل من الأسلوبين معاً ، ثم تنظم الكلمات والعبارات بعد ذلك وفق أوجه أو مجموعات متجانسة ، حيث يتم تحديد ما بين المصطلحات من علاقات . أما الطرق الآلية فتقدم تجميعات منظمة من المصطلحات التي غالباً ما تستخلص من الإنتاج الفكري . وتنطوي هذه الطرق على استخدام العديد من الأساليب الإحصائية واللغوية معاً في مختلف المراحل التي تمر بها العملية .

وغالباً ما تشمل إجراءات الاختيار الأساليب الإحصائية التي سبق أن عرضنا لها، والتي تعمل على تنحية الكلمات المستبعدة واختيار كلمات المحتوى (بعد تجريدها عادة وردها إلى جذوعها) والتي تتمتع «بقوة بيانية» جيدة (Luhn, 1957) أو «قيمة

تمييزية» (Salton, 1968). ويحدث في قليل من الأحيان استعمال الفئة النحوية (الفاعل ، المفعول) في اختيار المصطلحات (Hirschman, Grishman , and Sager, 1975). كما يتم أيضا اقتطاف العبارات الأكثر تخصيصاً والعبارات المعقدة اعتماداً على نوع معين من التحليل اللغوي ، وهو تحليل لا غنى عنه لضمان سلامة التتابع النحوي الناتج (Dillon and McDonald, 1983, Salton, 1988). كذلك تناول كل من لويس وكروفت (Lewis and Croft (1990) بالدراسة إجراءات التجميع العنقودي اعتماداً على العبارات اللغوية .

أما إجراءات التنظيم فإنها غالباً ما تجمع الكلمات أو العبارات إحصائياً أو تبعاً لتردداتها النسبي بصحبة بعضها البعض (Sparck-Jones, 1971) ، أو تبعاً لتردداتها النسبي بصحبة بعضها البعض (Stiles, 1961; Doyle, 1961). ويوضح الشكل رقم (٧٠) نتائج أساليب التجميع العنقودي الإحصائية . فمن الممكن أن تكون هناك مجموعات من المصطلحات (الكلمات أو العبارات) التي تعد مرتبطة ببعضها البعض دلاليًا (القسم A في الشكل رقم (٧٠) (Sparck-Jones, 1971; Van Rijisbergen, 1979) أو خريطة للمصطلحات ومدى قوة ما بينها من ارتباطات (القسم B في الشكل رقم (٧٠) (Doyle, 1961; Stiles, 1961). إلا أن ما لم تتعرض له هذه الأساليب هو الإعراب أو التعبير عما بين المصطلحات من علاقات دلالية . وترى سبارك جونز Sparck-Jones (1965) أن التجمعات العنقودية الناتجة على هذا النحو «غامضة دلاليًا» . هذا بالإضافة إلى أنه نظرًا لاستخدام النص الوارد في مجموعة بعينها من الوثائق في تكوين هذه التجمعات العنقودية ، فإن مدى تمثيل هذه التجمعات لمجال معين يمكن أن يكون محل تساؤل . ونظرًا لأنها تركز على الخواص الإحصائية للمجموعة التي تكونت منها ، فإن هذه التجمعات العنقودية قد لا تكون مفيدة بالنسبة للمجموعات الأخرى ، وربما أيضا بالنسبة لإحدى المجموعات في المجال الموضوعي نفسه.

وقد حظيت هذه القضايا بالاهتمام إلى حد ما من جانب فوكس وزملائه Fox et al. (1988) الذين يحاولون بناء مكنز (على هيئة شبكة دلالية) اعتماداً على معجم مسجل في شكل قابل للقراءة بواسطة الآلات . ويشتمل الشكل رقم (٧١) على مثال مبسط للنتائج التي انتهت إليها فوكس وزملاؤه . ومن المفترض أن اللغة المقننة الناشئة على هذا النحو لا يمكن أن ترتبط بمجموعة بعينها ، ومن ثم فإنها يمكن أن تمثل جوهر لغة عامة قابلة للاستخدام في كثير من الهيئات المختلفة.



الشكل رقم (٧٠) : نتائج التجميع الإحصائي للكلمات.

ويمكن للمكانز التي يتم بناؤها اعتماداً على مصادر عامة للمفردات كالمعاجم أن تكون مفيدة جداً، نظراً لإمكان تكوين مجموعات من الكلمات والعبارات المتصلة ببعضها البعض دلاليًا على نحو معين (كالجنس والنوع، والجزء والكل، والسبب والنتيجة أو الأثر مثلاً). إلا أن الأمر يتطلب جهداً حاسبياً ضخماً لبناء مثل هذه الشبكات، ولم يتضح بعد مدى جدوى المكانز التي يتم بناؤها اعتماداً على المعاجم القابلة للقراءة بواسطة الآلات، في عمليات الاسترجاع الفعلية، نظراً لأنها عادة ما تستند إلى المعاجم الإنجليزية المعيارية، ومن ثم فإنها تفتقر إلى كثير من المصطلحات المتخصصة التي تدعو الحاجة إليها بالنسبة لكثير من طلبات البحث الموضوعي المتخصص.

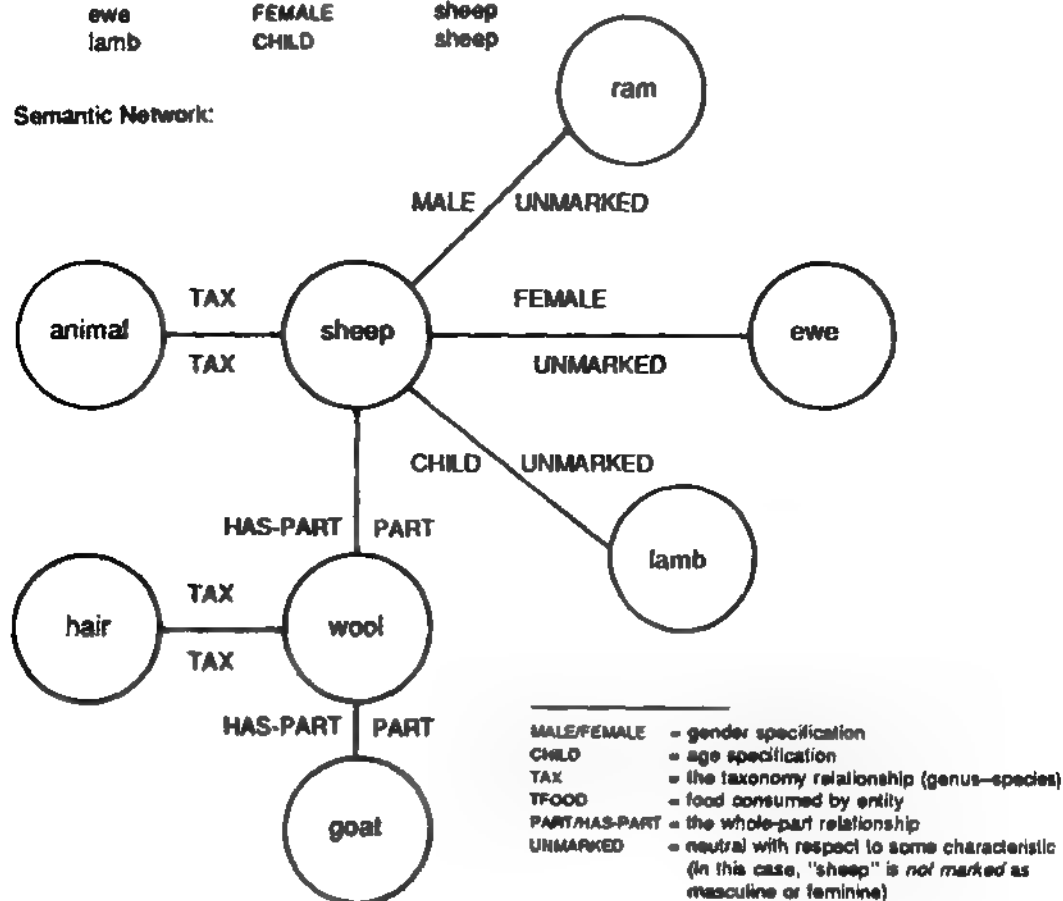
Dictionary Definitions

| | |
|------------------------|---|
| sheep n (pl unchanged) | grass-eating animal kept for its flesh as food (mutton) and its wool. |
| wool n [U] 1 | soft hair of sheep, goats, and some other animals . . . |
| ram n 1 | uncastrated male sheep. |
| ewe n | female sheep. |
| lamb n 1 [C] | young of the sheep . . . |

Extracted Lexical Relations:

| | | |
|-------|--------|--------|
| sheep | TAX | animal |
| sheep | TFOOD | grass |
| wool | PART | sheep |
| wool | PART | goat |
| wool | TAX | hair |
| ram | MALE | sheep |
| ewe | FEMALE | sheep |
| lamb | CHILD | sheep |

Semantic Network:



الشكل رقم (٧١) : نتائج التجميع اللغوي للكلمات والعبارات .

وهناك قضية أخرى تتصل بالتنظيم الإحصائي ، وهي أن التجمعات العنقودية التي يتم تكوينها عادة ، وكما سبق أن بينا ، لا تشتمل على المصطلحات المتصلة ببعضها البعض لغوياً (Lesk, 1969) . وقد طوّر باحثون آخرون طرقاً لا تنظم وفقاً لأسس لغوية ،

ومن بين هذه الطرق استخدام العلاقات المستقاة من المعاجم القابلة للقراءة بواسطة الآلات (Fox et al., 1988) والتحقق آلياً مما بين المصطلحات من علاقات اعتماداً على النصوص الطبيعية للوثائق (Warner, 1990; Ruge, 1992). وكما هو الحال بالنسبة لعمليات الاسترجاع الآلي الأخرى، فإنه ربما كان من الممكن للإجراءات الإحصائية واللغوية أن تفيد في مواقف مختلفة. وقد تبين لقلة من الباحثين فعلاً أن استعمال مجموعات المصطلحات المستندة إلى أسس لغوية لتوسعة الاستفسارات يؤدي في بعض الأحيان إلى الارتفاع بمستوى كفاءة الاسترجاع (Wang, Vandendorpe, and Evens, 1985).

التكشيف الآلي :

يقارن الشكل رقم (٧٢) بين الطرق المستخدمة في الأساليب اليدوية والآلية للتكشيف. وللتكشيف أهميته نظراً لأن القائمة الصغيرة بالمصطلحات المفتاحية أو الواصفات التي يتم اقتطافها أو تعيينها تعد معبرة إلى حد بعيد عن محتوى الوثيقة. وكما بينا في الفصل العاشر، فإن البحث بناءً على المصطلحات الكشفية غالباً ما يسفر عن معدل تحقيق أعلى مما يمكن الحصول عليه اعتماداً على اللغة الطبيعية. كذلك يمكن للتكشيف أن ييسر أيضاً الاستدعاء، نظراً لأن استعمال المصطلحات المخصصة التي يتم تعيينها من إحدى اللغات المقيدة، يساعد على ضمان تجميع الوثائق المتصلة بموضوعات معينة تحت شكل لغوي واحد.

وفي النظم اليدوية، فإن المصطلحات الكشفية عادة ما توزن تبعاً لمدى الاهتمام أو التركيز على كل عنصر من عناصر المحتوى الموضوعي للوثيقة. هذا بالإضافة إلى أن النظم التي تنطوي على روابط أو أدوار معينة كانت تستخدم في الستينيات وبداية سبعينيات القرن العشرين. فقد كان من المسلم به أن الأدوار تساعد على الحد من عدد الارتباطات الزائدة في نظام الاسترجاع، حيث تربط معاً المصطلحات المتصلة ببعضها البعض في الوثيقة، وتنحي ما عداها جانباً. أما الأدوار وهي أداة أكثر تعقداً وتطوراً فقد قطعت شوطاً أبعد في التعبير عما بين المصطلحات الكشفية من علاقات دلالية. (وقد تعرضنا للروابط والأدوار بمزيد من التفصيل في الفصل السادس).

ما يتم التعبير عنه : محتوى الوثائق

| أليستا | | | | يدويستا | | |
|---|--|--|---|---|---|-----------------------------------|
| تعين | | اقتطاف | | تعين | اقتطاف | |
| لغويا | إحصائيا | لغويا | إحصائيا | | | |
| من المكثز مع العلاقات الدلالية | من التجميعات المنقودية الآلية للكلمات | العبارات الاسمية | كلمات مفتاحية / ثنائيات كلمات | الترجمة الى واصفات لغة مقيدة | اختيار كلمات أو عبارات | إجرائات الاختيار |
| الأدوار والروابط الخاصة باصطلاحات المعينة | أوزان المصطلحات المعينة | العلاقات والأدوار الخاصة بالكلمات / العبارات | أوزان الكلمات / العبارات | الروابط والأدوار ، وأوزان الواصفات المخصصة أو غير المخصصة | الروابط ، والأدوار وأوزان الكلمات / والعبارات المخصصة وغير المخصصة | بيان العلاقات ذات الأهمية النسبية |
| الروابط أو الأدوار الخاصة باصطلاحات المعينة | المصطلحات المعينة الموزونة وغير الموزونة | الروابط والأدوار الخاصة بالكلمات / العبارات | الكلمات / العبارات الموزونة أو غير الموزونة | قوائم بالواصفات بالأوزان أو بالروابط أو الأدوار أو بدونها | قوائم بالكلمات / العبارات بالأوزان أو بالروابط أو الأدوار أو بدونها | شكل التعبير |

وعادة ما تكون أشكال التعبير الناتجة قوائم بالكلمات أو العبارات التي يتم اقتطافها من الوثائق أو يتم تعيينها من إحدى اللغات المقيدة. ويمكن للمصطلحات أن تكون في شكل قوائم بسيطة دون أية وسيلة لمزيد من التمييز ، كما يمكن أن تحدد لها أوزان أو روابط أو أدوار.

ويشتمل الإنتاج الفكري على وصف لبعض العمليات المناظرة في الطرق الآلية . ويمكن للعناصر أن تكون بعض كلمات المحتوى (مجردة في العادة) التي يتم اقتطافها من الوثائق (Salton , 1968) . هذا بالإضافة إلى أنه عندما تكون الكلمات المفردة عاجزة عن أن تكون وسائل تمييز جيدة ، فإنها أحياناً ما تضم في ثنائيات مع الكلمات التي تتردد بصحبتها ، في محاولة لتكوين ثنائيات أكثر تخصيصاً من الكلمات التي تتردد بمعدلات مقبولة (Salton, 1968) .

ولا ينطوي الأسلوب الذي عرضنا له تَوْأ على أية محاولة للتأكد من أن ثنائيات الكلمات تشكل عبارات نحوية صحيحة . إلا أن هناك أساليب أخرى لا تستخدم إلا العبارات النحوية فقط كمصطلحات كشفية ، ويتعين التقاط هذه العبارات من النص بواسطة نوع ما من التحليل اللغوي ، يتراوح بين النظم البسيطة التي تميز أنواع الكلمات ثم تتحقق بعد ذلك من العبارات الاسمية (Dillon and Gray, 1983) من جهة ، والإعراب الكامل للمدخلات والأكثر تعقداً من جهة أخرى (Fagan, 1988) .

والتكشيف الآلي بالتعيين أقل تداولاً وأكثر صعوبة من التكشيف الآلي بالاقتطاف أو الاشتقاق . وتطبيقاً لأحد الأساليب يتم اقتطاف المصطلحات إحصائياً كما سبق أن أوضحنا ثم تضاهى هذه المصطلحات بعد ذلك بالمصطلحات الواردة في التجمعات العنقودية المختزنة (كتلك التي تنتجها الطرق التي تناولتها سبارك - جونز Sparck- Jones, 1971 بالوصف) أو مقابل المصطلحات الواردة في الشكل الإلكتروني للمكانز اليدوية (Field, 1975) ، كما يتم تعيين المزيد من المصطلحات من المجموعات التي تمت مضاهاتها . وهناك طريقة أكثر تعقداً تقوم بإجراء تحليل لغوي للجمل الواردة في المدخلات حيث تختزل الجمل في صيغ بنيوية موحدة ، لتحويلها بشكل مطرد وناجح إلى مكنز يتخذ شكل الخريطة أو الشبكة (Evans et al., 1991) .

أما مدى التركيز النسبي على الجوانب الموضوعية التي تشملها الوثيقة فيتم التعبير عنه إحصائياً بوزن المصطلحات (Salton, 1968) ، في حين يتم التعبير عما بين المصطلحات من علاقات دلالية لغوياً بالروابط أو الأدوار (Courrier, 1980) . ويوضح الشكل رقم (٧٣) تنوع النتائج التي يمكن الحصول عليها من الكشف الآلي ، والتي تتراوح بين الكلمات المفتاحية أو العبارات الموزونة (A: Salton, 1968) ، والعبارات المركبة التي يتم تعيينها من أحد المكانز (B: Evans et al. 1991) إلى نظام سينتول SYN-TOL الأكثر تعقداً وتطوراً بمجموعاته من العلاقات الدلالية (C. Courrier, 1980) .

وتستخدم الغالبية العظمى من الجهود في الكشف الآلي الأساليب الإحصائية ، على أساس التسليم بأنه ما دامت الكلمة أداة تمييز جيدة فإنها يمكن أن تستخدم حينئذ كمصطلح كاشفي ، نظراً لأنها تفيد في بناء مجموعات أو فئات مناسبة من الوثائق . إلا أن الكشف البشري لا يمارس هذا الضرب من التمييز بوعي عند تحديد ما تدور حوله الوثيقة ، أو عند اقتطاف أو تعيين المصطلحات المناسبة . والأمل معقود على إمكان أن تسفر الأساليب الإحصائية عن مجموعات مقبولة من المصطلحات الكشفية ، على الرغم من أن المصطلحات عادة ما يتم تحديدها أو التوصل إليها على نحو مختلف في الطرق الآلية.

ومما لا شك فيه ، أن من بين أسباب استخدام الأساليب الإحصائية في الكشف الآلي (وهو أمر له ما يبرره) صعوبة أتمتة القواعد التي يتبعها الكشف البشري . ومن بين الأساليب الأخرى التخلي عن مبدأ الكشف الآلي الكامل واتباع بعض الأساليب التي تستعين بالآلات لأتمتة الجوانب الأيسر من غيرها في عملية الكشف ، وإفساح مجال المشاركة من جانب الكشف في مختلف المراحل حسبما تقضي الضرورة.

ومن بين الجهود في هذا الصدد ما قدمه كلنجبيل (Klingbiel 1973) باستعمال معاجم الكلمات وسلاسل الكلمات المتاحة لتمييز واقتطاف المصطلحات الكشفية المرشحة من النصوص الطبيعية ، حيث يتم تقييم هذه المصطلحات وقبولها أو رفضها من جانب مكشف بشري . ويوضح الشكل رقم (٧٤) عينة من مخرجات هذا النظام وهي عبارة عن اقتباس مجزأ ، بالإضافة إلى المصطلحات الكشفية المرشحة التي تحقق منها أو حدها.

A.

* TEXT 2MICRO-PROGRAMMING .

\$MICRO-PROGRAMMING

\$R. J. MERCER (UNIVERSITY OF CALIFORNIA)

\$U.S. GOV. RES. REPTS. VOL 30 PP 71-72(A) (AUGUST 15,1958) PB 126893

MICRO-PROGRAMMING . THE MICRO-PROGRAMMING TECHNIQUE OF DESIGNING THE CONTROL CIRCUITS OF AN ELECTRONIC DIGITAL COMPUTER TO FORMALLY INTERPRET AND EXECUTE A GIVEN SET OF MACHINE OPERATIONS AS AN EQUIVALENT SET OF SEQUENCES OF ELEMENTARY OPERATIONS THAT CAN BE EXECUTED IN ONE PULSE TIME IS DESCRIBED .

| | | | | | | | | | | |
|--------------|--------|----|--------|----|--------|----|--------|----|--------|----|
| 2MICRO-PROGR | CIRCUI | 12 | COMPUT | 12 | CONTRO | 12 | DESCRI | 12 | DESIGN | 12 |
| | DIGIT | 12 | ELECTR | 12 | ELEMEN | 12 | EQUIVA | 12 | EXECUT | 24 |
| | FORM | 12 | GIVE | 12 | INTERP | 12 | MACHIN | 12 | OPERAT | 24 |
| | PULSE | 12 | SEQU | 12 | SET | 24 | TECHNI | 12 | TIME | 12 |

B.

Title: "A data/knowledge base management testbed and experimental results on data/knowledge base query and update processing."

Presents experience in designing and implementing a data/knowledge base management testbed. The testbed consists of two layers, the knowledge manager and the database management system, with the former at the top. The testbed is based on the logic programming paradigm, wherein data, knowledge, and queries are all expressed as Horn clauses. The knowledge manager compiles pure, function-free Horn clause queries into embedded-SQL programs, which are executed by the database management system to produce the query results. The database management system is a commercial relational database system and provides storage for both rules and facts. First, the testbed architecture and major data structures and algorithms are described. Then, several preliminary tests conducted using the current version of the testbed and the conclusions from the test results are presented. The principle contributions of this work have been to unify various concepts, both previously published and new ones developed, into a real system and to present several insights into data/knowledge base management system design gleaned from the test results and the design and implementation experience.

• Exact Matches to Certified Terms in the Thesaurus:

- HORN-CLAUSE
- ALGORITHM
- DATABASE MANAGEMENT SYSTEM

الشكل رقم (٧٣) : نتائج الكشف الآلي .

• Matches **to** More General Terms in the Thesaurus:

- _ DATA/KNOWLEDGE \leftarrow Date/Knowledge Base Management Testbed
- _ DATABASE \leftarrow Database Management System
- _ RELATIONAL-DATABASE \leftarrow Commercial Relational Database System
- _ DATA-STRUCTURE \leftarrow Major Data Structures and Algorithms
- _ DATABASE MANAGEMENT \leftarrow Database Management System
- _ LOGIC PROGRAMMING \leftarrow Logic Programming Paradigm
- _ ARCHITECTURE \leftarrow Testbed Architecture
- _ LOGIC \leftarrow Logic Programming Paradigm

• Novel Terms (in the Text) that Utilize Concepts Represented in the above Sets of Terms (hence are Specified Matches) or have "high" Compositional or Scored Matches:

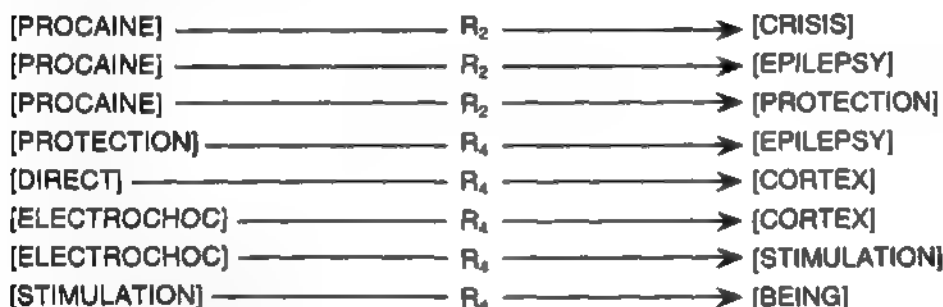
- _ FUNCTION-FREE HORN-CLAUSE QUERY
- _ DATA/KNOWLEDGE BASE QUERY
- _ DATA/KNOWLEDGE BASE MANAGEMENT TESTBED
- _ TEST ARCHITECTURE
- _ QUERY RESULT
- _ MAJOR DATA-STRUCTURE
- _ DATA/KNOWLEDGE BASE MANAGEMENT SYSTEM DESIGN
- _ EMBEDDED-SQL PROGRAM
- _ COMMERCIAL RELATIONAL-DATABASE SYSTEM
- _ LOGIC PROGRAMMING PARADIGM

C.

Abstract in natural language:

Experimental observations on the prevention of seizures by intravenous procaine injection. Ten monkeys were stimulated by a direct cortical electrochoc. An intravenous injection of procaine has protected them from epilepsy for 30 min.

SYNTOL representation:



R_2 : Associative relation

R_4 : Coordinative relation (the coordinated relation is not oriented)

Evaluation problems in interactive information retrieval - G. Salton: Department of Computer Science, Cornell University, Ithaca, New York 14850

1 2 6 1 1 1 1 1 1 1 1 1 6
Summary - Interactive retrieval procedures are normally based on rapidly accessible files. Special storage
6 4 1 6 6 1 1 4 1 6 6 1 1 1 1 1 1
organizations and file search techniques are used, and the system user is made to fulfill an important role
1 1 6 1
during the retrieval process.
1 1 1 1 1 2 6 6 1 1 1 1 1 6 1
In the present study, the interactive retrieval environment is briefly examined. The special problems which
1 1 1 1 8 2 6 1 1 1 4 1 1 1 1 1
arise in the evaluation of interactive retrieval are then discussed, and methods are described for evaluating
2 1 1 4 6 6 6 1 1 1 1 1 3 6
partial file searches and user feedback techniques. Evaluation results obtained with the SMART system
1 1
are presented. (Inform. Stor. Retr. 1970, 6, 29 - 44.)

المصطلحات الكشفية المرشحة

| | |
|-----------------------------------|----------------------------|
| INTERACTIVE RETRIEVAL | (2,6--adj.,noun) |
| STORAGE ORGANIZATIONS | (6,6--noun,noun) |
| SEARCH TECHNIQUES | (6,6--noun,noun) |
| SYSTEM USER | (6,6--noun,noun) |
| INTERACTIVE RETRIEVAL ENVIRONMENT | (2,6,6--adj.,noun,noun) |
| INTERACTIVE RETRIEVAL | (2,6--adj.,noun,) |
| USER FEEDBACK TECHNIQUES | (6,6,6--noun,noun,noun) |
| SMART SYSTEM | (3,6--proper noun, noun) |

الشكل رقم (٧٤) : مثال لمخرجات إحدى عمليات التكشيف التي يتم فيها الاستعانة بالآلات.

وهناك مشروع آخر أحدث وأكثر طموحًا ، في المكتبة القومية للطب ، يعتمد على رصيد معرفي يتكون من إطارات لتوجيه المكشفين نحو فئات بعينها من المصطلحات الكشفية ، ولتقديم مدخلات - آليا - تعتمد على المعلومات التي أمكن الحصول عليها باستخدام استنتاجات تستثمر العلاقات التي تم التعبير عنها رمزيًا بوضوح والواردة في الإطارات (Humphrey and Miller, 1987). ونجد في الشكل رقم (٧٥) نموذجًا مبسطًا لمخرجات هذا النظام.

صياغة الاستفسارات آليا :

إذا سلمنا بأن الوثائق والاستفسارات ينبغي أن تضاهى عند مستوى معين لكي يكون الاسترجاع ناجحًا ، فإنه لا عجب أن يكون هناك كثير من مظاهر التشابه بين وسائل التعبير عن خصائص كل منهما أو تمثيلهما في كل من النظم اليدوية والنظم الآلية. وكما يتبين من الشكل رقم (٧٦) فإن الاستفسارات يمكن أن تتكون من كلمات وعبارات باللغة الطبيعية المستمدة من اللغة الخاصة بمسئول البحث ، أو تلك التي يمكن أن تقتطف من طلب البحث الذي يقدمه المستفيد . كذلك يمكن لمسئول البحث أن يستعمل

Indexer prompted for the following slot values:

disease : Cysts

body-part: value: Popliteal Artery

procedure: value: Angiography + Tomography X-Ray Computed
+ Ultrasonic Diagnosis

symptom: value: Intermittent Claudication

Program substitutes legal Medical Subject Headings [words in bold face indicate substituted descriptors]

(setq frame:

```
( cyst _85140959 (is-a (value cyst))  
  (body-part (value popliteal_artery))  
  (procedure  
    (value angiography  
      x-ray_computed_tomography  
      ultrasonography))  
  (symptom (value Intermittent_claudication))  
  (contained-in (value document_85140958))))
```

الشكل رقم (٧٥) : نموذج من مخرجات نظام للتكشيف يستعين بالآلات في المكتبة القومية للطب.

الواصفات المعينة من إحدى اللغات المقيدة بدلا من مصطلحات البحث المعتمدة على اللغة الطبيعية أو بالإضافة إلى هذه المصطلحات. ويتعين في كثير من النظم الربط بين المصطلحات بقواعد الجبر البوليني كما سبق أن بينا في الفصل الثامن. كذلك يمكن أيضًا وزن المصطلحات تبعًا لما لها من أهمية في الطلب. كما يحدث في أحيان نادرة أن يتم الربط بينها بواسطة الروابط والأدوار. ولا يمكن بالطبع استعمال الوزن والروابط والأدوار في الاستفسارات إلا إذا كانت قد استعملت أيضًا في التعبير عن المحتوى الموضوعي للوثائق. وعادة ما تكون النتيجة هي تحويل طلب البحث الذي يقدمه المستفيد مصاغًا باللغة الطبيعية في العادة، يدويًا، إلى سلسلة من المصطلحات الكشفية، المرتبطة ببعضها البعض في غالب الأحيان، بقواعد الجبر البوليني، وأحياناً ما يتم وزن هذه المصطلحات أو ربطها ببعضها البعض دلاليًا.

وتبدأ الطرق المعتمدة على الآلية اعتماداً كاملاً بصياغة الاستفسار باللغة الطبيعية، حيث يتم تجهيز هذه الصيغ لاقتطاف المصطلحات الضرورية فضلاً عن العوامل المنطقية والأوزان والروابط والأدوار. والهدف في جميع الحالات هو إنتاج صيغة استفسار بالعناصر نفسها كما هو الحال بالنسبة للوثائق في النظام. ونجد في الشكلين رقمي (٦٨ و ٧٧) نماذج للصيغ الآلية للاستفسارات.

ما يتم التعبير عنه : مضمون الاستفسار

| آلياً | | | | يدوياً | | |
|---|---------------------------------------|--|--|--|--|-----------------------------------|
| تعيين | | اقتطاف | | تعيين | اقتطاف | |
| لغويًا | إحصائيًا | لغويًا | إحصائيًا | | | |
| من الكائنات مع العلاقات الدلالية | التوسع بعناقيد الكلمات/ العبارات | العبارات الاسمية نحويًا | كلمات، والـ / ثنائيات كلمات | الترجمة الى واصفات اللغة المقيدة | انتقاء كلمات/ عبارات من الاستفسار | إجراءات الانتقاء |
| العوامل البوليئية، الروابط، الأدوار بين الكلمات/ العبارات | أوزان الكلمات/ العبارات المعينة | العوامل البوليئية، والروابط والأدوار بين الكلمات/ العبارات | أوزان الكلمات/ العبارات | العوامل البوليئية، الروابط، الأدوار وتحديد الأوزان | العوامل البوليئية، الروابط، الأدوار وتحديد الأوزان | بيان العلاقات ذات الأهمية النسبية |
| العوامل البوليئية، الروابط، الأدوار بين الكلمات/ العبارات | المصطلحات المعينة/ موزونة/ غير موزونة | العوامل البوليئية، الروابط، الأدوار بين الكلمات/ العبارات | كلمات/ ثنائيات كلمات، موزونة/ غير موزونة | أوزان ، عوامل بوليئية، روابط، أدوار خاصة بالواصفات المعينة | أوزان ، عوامل بوليئية، روابط، أدوار خاصة بالكلمات/ العبارات المنتقاة | شكل التعبير |

الشكل رقم (٧٦) : مقارنة بين الأساليب اليدوية والأساليب الآلية لتحليل الاستفسارات والتعبير عنها.

وفي الأساليب الإحصائية يتم اقتطاف الجذور الرئيسة وثنائيات الجذور مباشرة من الصيغة التي عبر بها المستفيد (مع تجاهل الكلمات المستبعدة وغيرها من العبارات غير الجوهرية) حيث يتم وزن هذه الجذور ثم مقارنتها بما ورد في تسجيلات الوثائق كما في الشكل رقم (٦٨) (Salton, 1968) . هذا بالإضافة إلى أنه من الممكن تقديم الوثائق المسترجعة على هذا النحو إلى المستفيد للحكم على مدى صلاحيتها ، كما يمكن تعديل الاستفسار وإعادة تقديمه إلى النظام، بطريقة تكرارية مستمرة تعرف « بالتلقيم المرتد الخاص بالصلاحية » (Salton, 1968) . كذلك يمكن للعناصر المستمدة من صيغة البحث أن يتم الربط بينها بالقواعد البولينية ، إما عن طريق التجهيز اللغوي للاستفسار لتحديد كيفية الربط بين العناصر منطقياً (Das Gupta, 1987) أو إحصائياً (حيث يتم حساب عدد الوثائق المسترجعة لكل صيغة ربط بولينية ، والتوصل إلى صيغة بناء على عدد محدد معين للوثائق المسترجعة) (Salton, Buckley and Fox, 1983) .

وبدلاً من ذلك يمكن إعراب المدخلات جزئياً أو كلياً للعثور على العبارات الاسمية نحويّاً (Dillon and Gray, 1983; Croft, Turtle and Lewis, 1991) . ويحدث في أحيان قليلة أن يتم تعيين المصطلحات من مكنز ناتج عن التجمعات العنقودية بناء على الأساليب الإحصائية التي تعرضنا لها في القسم الخاص ببناء المكنز آلياً ، (Sparck-Jones, 1971) (أو يدوياً ، وبذلك يمكن إضافة المصطلحات في علاقات دلالية محددة إلى الاستراتيجية . وعادة ما يكون من الممكن التحقق من الأدوار أو الروابط الدلالية أو النظمية ، بإجراء المزيد من عمليات التجهيز للمدخلات، للعثور على المفاتيح اللفظية التي تدل على مختلف العلاقات الدلالية كالعامل أو السبب في مقابل النتيجة أو الأثر، أو الجزء في مقابل الكل (Courier, 1980) .

وفي مقابل الطريقة التي يتم فيها تصميم عمليات تجهيز الوثائق والاستفسارات كجزء من نظام تجريبي كامل، هناك طرق أخرى يقصد منها العمل « كواجهات تعامل » أو « منافذ » للنظم القائمة، حيث تم تحليل الوثائق وتجهيزها لأحد النظم العاملة فعلاً والمتاحة تجارياً . وتهدف هذه النظم لأن تجعل من تعقيدات عملية البحث نظاماً متطوراً واضحاً بالنسبة للمستفيد النهائي . ومن ثم فإنه يتم التركيز في وضع أدوات البحث هذه على تحديد تلك الوظائف التي ينبغي أن تتضمنها لمساعدة المستفيد على أن يعرب،

A. Automatic Boolean query formulation (compared with manual):

| <i>Manual Form (Boolean Statement)</i> | <i>Automatic Form</i> |
|---|--|
| 1. (catalogue or catalog) and (mechanization or automatic or computerization) | [(<catalogue, 6.36> or (2) <catalog, 5.36>), 5.86] and 1.5 [mechanization, 4.04] and (1.5) [<automation, 2.66> and (1.5)] [(<computerization, 1.60>), 2.13]. |
| 2. (information and science) and (education or training) | [<information, 0.90>] and (1.5) [<science, 2.19>] and (1.5) [(<education, 4.36> or (1.5) <training, 3.78>), 4.06] |

**B. Search statement: Automatic manufacture of mesh screens for widgets
[semantic processor]**

Two groups of searchable paraphrases:

| | |
|---------------------------|-----------------------------|
| AUTOMATIC MANUFACTURE | SCREENS FOR WIDGETS |
| AUTOMATED MANUFACTURE | SCREENS OF MESH FOR WIDGETS |
| AUTOMATIC MANUFACTURING | MESH SCREENING FOR WIDGETS |
| MANUFACTURE BY AUTOMATION | MESH SCREENING FOR A WIDGET |

C. Search statement: Side effects of aspirin on the liver

[yields]

[ASPIRIN] _____ R₂ _____ [LIVER]

agent

consecutive

patient

relation

الشكل رقم (٧٧) : نماذج للصيغ الآلية للاستفسار.

وبشكل كامل عن عناصر البحث، وذلك بتوفير أو اقتراح الأشكال المختلفة المناسبة لمصطلحات البحث، فضلاً عن أدوات الربط المنطقية أو اللغوية. والسبيل المألوف للتفاعل مع المستفيد في هذه النظم، هو وضع برنامج يمكن أن يكون بمثابة نظام إرشادي آلي للبحث، أو كما يسمى في غالب الأحيان « نظام وسيط خبير ». وهناك العديد من مثل هذه النظم الآن، بما في ذلك كانسيرش CANSEARCH (Pollitt, 1987)، وهو نظام خبير للبحث في الإنتاج الفكري لعلاج السرطان، وبلكسوس PLEXUS (Vickery and Brooks, 1987) وهو نظام خبير للبحث في مرصد بيانات مرجعي في فلاحه البساتين. ويصف فيكري (Vickery (1992 واجهات التعامل الذكية المتاحة حالياً في أوروبا، كما يقدم باروت (Parrott (1992 وصفاً تحليلياً لبعض الواجهات الذكية التي طُورت في الولايات المتحدة. ومعظم هذه الواجهات تسترجع الإشارات الوراقية، وقد أدت الجهود التي بذلت مؤخراً إلى تطوير واجهات خاصة بنظم النصوص الكاملة (Gauch and Smith, 1993).

وتصف بروكس (Brooks (1987 القدرات العامة التي تميز النظام الخبير؛ فالنظام ينبغي أن يكون قادراً على تحقيق مستوى الكفاءة نفسه الذي يحققه الخبير البشري في المهمة نفسها، كما ينبغي أن تكون المشكلة التي صمم النظام للتصدي لها معقدة إلى الحد الذي تتطلب معه ذكاء الخبير البشري لحلها. كذلك ينبغي أن يكون النظام الخبير قادراً على التعامل مع المشكلة التي يتم التعبير عنها بشكل أولي كيفما اتفق، وترجمتها إلى صورة داخلية يمكن تجهيزها اعتماداً على القواعد الخبروية للنظام وأرصدة معارفه. وأخيراً ينبغي للنظام أن يكون قادراً على استنباط ما يقوم به من عمليات التجهيز وتفسير قراراته للمستفيد.

وللاضطلاع بهذه الوظائف فإن وسيط البحث الخبير قد يحتاج إلى المكونات التالية:

١ - المعرفة : فالمعرفة التي يحتاج إليها مثل هذا النظام شاملة، وتضم المعرفة الخاصة بالمهمة (قواعد استراتيجية البحث، والتفاوض والصياغة) والمعرفة الخاصة بما بين الوثائق وتسجيلاتها من علاقات، والمعرفة الخاصة بقواعد التعامل مع مرادف البيانات (أي قواعد صياغة استراتيجية البحث) والمعرفة الخاصة بالمستفيد وما يتصل منها بالحاجة إلى المعلومات على وجه التخصيص، وبالمجال الموضوعي ومصطلحاته.

٢ - الاستنباط : وينبغي أن تشتمل أساليب الاستنباط على إجراءات للاستدلال على استراتيجيات البحث الملائمة ، اعتماداً على الحقائق المستمدة من أرصدة المعارف التي أشرنا إليها توأ.

٣ - واجهة المستفيد : ينبغي للنظام ، عند صياغة الاستراتيجية ، أن يكون بمثابة المرشد، وأن يتفاعل مع المستفيد بطريقة تكافلية مرنة، تجعل النظام الأساس واضحاً بالنسبة للمستفيد. كما أنه ينبغي أن يكون قادراً على أن يوضح للمستفيد كيف توصل إلى ما انتهى إليه من نتائج.

وتطوير النظم الخبرة جهد مستمر ، بدأ في منتصف ثمانينيات القرن العشرين (Smith, 1987) . وتقارير مثل هذه النظم في الإنتاج الفكري، في واقع الأمر عبارة عن وصف لنماذج أولية مصممة هندسياً على أسس واعية، لا لنظم عاملة فعلاً. وتحد هذه النظم قيود لا يستهان بها ؛ قيود في مجالاتها الموضوعية ، وفي مدى ما يمكن أن تتعامل به من مصطلحات، وفي تنوع ما يمكن أن تتقبله من المستفيد من مدخلات، وفي مدى مرونة الاستجابة لأخطاء المستفيد . إلا أنها رغم ذلك قد أكدت نجاحها في المجالات الموضوعية الضيقة، كما ألفت ولا شك الضوء على ما يحتاج الباحثون إلى معرفته حول عملية البحث والسياق الذي تتم فيه ، لأجل بناء نظم أكثر كفاءة وفعالية.

الاستخلاص الآلي :

لا يمكن للطرق اليدوية للاستخلاص ، أن تستخدم كما هو موضح في الشكل رقم (٧٨) ناتج إعادة صياغة الوثائق أو ملخصاتها الحقيقية فحسب، وإنما يمكن أن تعتمد أيضاً على الجمل التي تقتطف مباشرة من الوثائق الأصلية ، وإن كان هذا الأسلوب غير مألوف. وغالباً ما يتم في الاستخلاص البشري تقديم نص يتناول أهداف المؤلف الأصلي ومناهجه ونتائجه والخلاصة التي انتهى إليها ، وفيما عدا ذلك فإن المستخلصات يمكن أن تتفاوت أسلوبياً فيما بينها تفاوتاً ملحوظاً .

أما في الطرق الآلية فالأقتطاف أكثر شيوعاً من الاستخلاص. وكما هو الحال في الكشف فإن الخطوة الإضافية الخاصة بتحويل النصوص المقتطفة إلى شكل جديد، تتسم بالصعوبة، كما أنه لا يمكن الحكم ما إذا كانت النتائج أفضل بكثير من طريقة الاقتطاف البسيطة.

ومن الممكن اقتطاف الجمل آلياً إما بالأساليب الإحصائية أو بالأساليب اللغوية. وكما هو الحال في التكشيف فإن بعض الجهود الرائدة في هذا المجال يرجع الفضل فيها إلى لون (1958) Luhn الذي قام بحساب رصيد إجمالي للجمل بعد استبعاد الكلمات الوظيفية العامة وتلك الكلمات التي تتردد بمعدل أقل من حد معين. وتستخدم الأساليب الإحصائية الأخرى تنويعات لهذا الأسلوب الأساس ، وتشمل هذه التنويعات إعطاء أرصدة مرتفعة للجمل التي تشتمل على عبارات (Luhn, 1958) ، والجمل التي تضاهي كلمات المضمون الواردة بها الكلمات التي ترد في العناوين أو العناوين الفرعية (Edmundson, 1969) أو الجمل التي ترد في بداية الفقرات أو في نهايتها (Edmundson, 1969) .

وتقوم بعض الأساليب الأخرى على حساب أرصدة رقمية للجمل ولكنها تستخدم معايير لغوية في العملية، حيث تتخذ قائمة الكلمات والتعبيرات التي تدل بوجه خاص على محتوى الوثيقة أو لا تدل على هذا المحتوى ، كأساس لإعطاء الجملة رصيذاً إيجابياً أو سلبياً. فكلمات مثل « الغرض » و « البحث الحالي » تعطي أوزاناً إيجابية ، بينما تعطي كلمات مثل « لا أهمية له » أو « من المستحيل » قيماً سلبية (Rush, Salvador, and Zamora, 1971 . ومن الممكن في بعض الأحيان تجهيز تعبيرات أكثر تعقداً مثل « إن الهدف الأساس للبحث الحالي هو وصف ... » أو « تهدف هذه المقالة إلى مراجعة ... » (Paice, 1990) .

وتنطوي إحدى الإضافات الأكثر تطوراً وتعقداً للاستخلاص الآلي على مراعاة بعض جوانب معينة في بنية النص، وخاصة تكرار الإشارة إلى كثير من الموضوعات أو الأفكار أو المفاهيم عدة مرات باستعمال مرادفات مختلفة، وكذلك « التعبيرات المحددة للعلاقة » كالضمائر . ويصف بلاك (1990) Black كيف يمكن للأساليب الإحصائية للاستخلاص الآلي أن تتحقق من المفاهيم وتحصيها ، وذلك باقتطاف وإحصاء جميع الكلمات التي تدل على المفهوم. فقد تكرر ذكر الفكرة التي يدل عليها البنط الأسود في الجملة التالية ، على سبيل المثال مرتين : “ As the ISO standards become more stable, they will replace the coloured book protocols”.

وفي مقابل إجراءات الاقتطاف هناك عدد قليل من نظم الاستخلاص الآلي الحقيقية، حيث يختلف النص عما كان عليه في الأصل. وقد قدم كل من ماثيس وراش ويانج

ما يتم التعبير عنه : محتوى الوثيقة

| آلية | | يدوي | | |
|---------------------------------|---|--|-------------------------------------|------------------|
| تخصيص | اقتطاف | تخصيص | اقتطاف | |
| تعديل الجمل بالتحليل اللغوي | اختيار الجمل وفقاً لأعلى رصيد رقمي | تخصيص المفاهيم / الجمل المفتاحية أو الدالة | اختيار الجمل الدالة أو المفتاحية | إجراءات الاختيار |
| تنظيم الجمل بالتحليل البنيوي | تنظيم الجمل وفقاً لتسلسل الترتيب إحصائياً أو وفقاً للتسلسل الذي ترد به | التنظيم في فقرة مترابطة | التنظيم في فقرة مترابطة | إجراءات التنظيم |
| مستخلص | مقتطف | مستخلص | مقتطف | شكل التعبير |

الشكل رقم (٧٨) : مقارنة بين الطرق اليدوية والطرق الآلية لإجراءات الاستخلاص وصياغة المستخلصات .

A. Keyword approach

Document

Source: The Scientific American, Vol. 196, No. 2, 86-94, February, 1957

Title: Messengers of the Nervous System

Author: Amodeo S. Marrazzi

Auto-Abstract

It seems reasonable to credit the single-celled organisms with a system of chemical communication by diffusion of stimulating substances through the cell, and these correspond to the chemical messengers (e.g., hormones) that carry stimuli from cell to cell in the more complex organisms (7.0)

Finally, in the vertebrate animals there are special glands (e.g., the adrenals) for producing chemical messengers, and the nervous and chemical communication systems are intertwined: for instance, release of adrenalin by the adrenal gland is subject to control both by nerve impulses and by chemicals brought to the gland by the blood (6.4)

The experiments clearly demonstrated that acetylcholine (and related substances) and adrenalin (and its relatives) exert opposing actions which maintain a balanced regulation of the transmission of nerve impulses (6.3)

It is reasonable to suppose that the tranquilizing drugs counteract the inhibitory effect of excessive adrenalin or serotonin or some related inhibitor in the human nervous system (7.3)

[Sentences were selected by statistical analysis when they had a degree of significance of 6 and over.]

B. Cue method plus grammatical modification

Document

Source: Journal of the American Society for Information Science; March-April, 1973: 101-110

Title: Improvement of Automatic Abstracts by the Use of Structural Analysis

Author: Mathis, Betty A.; Rush, James E.; Young, Carol E.

[auto-abstract of their article describing their method]

Abstract

We have undertaken to extend the capabilities of the abstracting system described by Rush, Salvador and Zamora by adding to the system a modification procedure that could be employed to make the abstracts produced by the system more acceptable to the reader. Results of this study are reported in this paper. Our purpose is to present a rationale for the modification phase of an abstracting system and to describe several modification rules whose implementation is an initial step toward the automated production of abstracts that contain sentences written especially for the abstract. We have described several methods for improving the readability of abstracts produced by computer program. The research described in this paper was performed as a part of a larger project whose aim is the development of an operational automatic abstracting system.

[Cue words: undertaken, purpose, methods, aim]

[Modification of text: We have therefore undertaken to extend the capabilities of the abstracting system described by Rush, Salvador and Zamora--deleted therefore]

(Mathis, Rush, and Young (1973) وصفاً لإضافة إلى أحد نظم الاقتطاف أمكن بمقتضاها تعديل الجمل وخاصة لصياغة المستخلص. وكان هذا الجهد ينطوي في الأساس على الربط بين الجمل لتشكل تعبيرات أكثر تعقداً أو أكثر إحكاماً . فمن الممكن على سبيل المثال للجملتين:

Individual manufacturers offer ALGOL, BASIC and FOCAL Compilers.

و

Most manufacturers offer programming support on an individually negotiated contract basis.

أن يتم الربط بينهما على النحو التالي :

Individual and most manufacturers offer ALGOL, BASIC and FOCAL Compilers, and programming support on an individually negotiated basis, respectively.

وبالإضافة إلى الحاجة إلى التحقق من الجمل واقتطافها ، هناك الحاجة إلى تنظيمها بطريقة ما. وهناك طريقة ترتب الجمل في المستخلص النهائي وفقاً لتناقص الوزن الإحصائي. كذلك أضافت الطرق الأحدث من ذلك إمكانية جمع كل الجمل التي تدل على المفهوم نفسه معاً، حيث يؤدي ذلك إلى تقديم مستخلصات تعد أكثر تماسكاً، لأنها لا تشتمل إلا على قليل من الإشارات غير المعروفة ، وقد لا تشتمل على مثل هذه الإشارات مطلقاً (Paice, 1990) .

ويوضح الشكل رقم (٧٩) النتائج التي أمكن الحصول عليها من نظامين آليين للاستخلاص. والقسم A ناتج عن أسلوب بسيط لإحصاء تردد الكلمات، أما القسم B فناتج عن إجراءات أكثر تعقداً، تبحث عن ورود الكلمات التي يمكن أن تقدم جملاً تدل على المحتوى بشكل أفضل من غيرها (« كفرض » و « هدف » على سبيل المثال) .

الطرق الآلية المستخدمة في الفهارس المتاحة على الخط :

كذلك أجريت بعض التجارب لاستخدام الإجراءات الآلية أو شبه الآلية، في فهارس المكتبات المتاحة على الخط المباشر ، وإن كان البحث في هذا المجال ليس بكثافة أو اطراد البحث في الملفات الوراقية الأخرى المتاحة على الخط المباشر. ومن بين أسباب ذلك أن

الفهارس المتاحة على الخط المباشر لم تستخدم على نطاق واسع إلا في بداية ثمانينيات القرن العشرين. وهناك سبب آخر وهو اعتقاد قديم بأن العزوف عن البحث بالمدخل الموضوعي في الفهارس الورقية كان يعبر عن الاحتياجات الحقيقية للمستفيدين بالنسبة للفهارس الإلكترونية أيضاً، وهو اعتقاد سرعان ما توارى نتيجة للدراسات الخاصة بالإفادة من الفهارس المتاحة على الخط المباشر. وقد أكدت هذه النتيجة، على نطاق واسع، الدراسة التي أجراها مجلس الموارد المكتبية على المستوى القومي (Matthews, Lawrence and Ferguson, 1983) والتي كشفت عن الاستخدام المكثف للبحث من جانب المستفيدين، وفقاً للمدخل الموضوعي في الفهارس التي تجاوزت مرحلة التجريب، وكذلك عن رغبة المستفيدين في توافر إمكانات البحث بالمدخل الموضوعي أكثر كثافة وأكثر مرونة. ولا عجب إذن أن تتضمن الدراسات التجريبية التي تتناول الفهارس المتاحة على الخط المباشر، إجراء اختبارات للكثير من مقومات البحث بالمدخل الموضوعي.

ومشروع أوكابي OKAPI Project (Walker, 1987 , 1988) عبارة عن فهرس تجريبي جارٍ على الخط المباشر بمعهد وسط لندن للتقانة Polytechnic of Central London ويتيح هذا النظام للمستفيدين إمكانية إدخال الكلمات أو العبارات التي يتم تجريدها فيما بعد والبحث عنها في جميع الحقول المصدرية المعبرة عن المحتوى الموضوعي، بما في ذلك العناوين والعناوين الفرعية ورءوس الموضوعات، وتبصرات المحتوى. فإذا وردت كلمتان أو أكثر في الكشف، فإنها ترتبط ببعضها البعض آلياً بأداة العطف البولينية «و». فإذا لم تكن هناك تسجيلات تضاهي الطلب وفقاً لهذه الاستراتيجية، يتم إعطاء الكلمات أوزاناً تتناسب عكسياً مع ترددها في الملف. ومن ثم فإن نظام أوكابي يعمل بكل من العوامل المنطقية والأوزان الإحصائية.

وقد استخدم الأسلوب الإحصائي، وربما بشكل أكثر كثافة في سايت CITE ، وهو فهرس تجريبي على الخط المباشر، تطور في المكتبة القومية للطب NLM (Doszkocs, 1983). وكان نظام سايت هذا قادراً على التعامل مع اللغة غير المقيدة، بما في ذلك الجمل أو الفقرات فضلاً عن الكلمات والعبارات. وكانت الكلمات يتم تجريدها، حيث يتم إعطاء العناصر الناتجة أوزاناً بناءً على عكس معدلات ترددها في مجموعة الوثائق. وكان من بين

الملاحم الخاصة بالنظام قدرته على البحث في قائمة رموس الموضوعات الطبية *Medical Subject Headings* ، واسترجاع واقتراح المصطلحات الكشفية المحتملة، التي كان من الممكن استعمالها حينئذ في البحث (Doszkocs, 1978) .

ويصف لارسون (1992) Larson فهرساً تجريبياً على الخط المباشر أحدث مما سبق، وهو تشيشاير CHESHIRE . ويعتمد هذا المشروع على شكل معدل من نظام سمات SMART الخاص بسالتون Salton ، لاختبار مدى فعالية عدد من أدوات الاسترجاع الآلي الشهيرة، وذلك في مجموعة اختبار قوامها حوالي ٣٠٠٠٠ من تسجيلات مارك MARC . وتشمل الطرق عدة نماذج مختلفة للاسترجاع (منها ما يعتمد على الكم الموجه Vector ومنها ما هو احتمالي) والتوسع في الاستفسارات بالتجمعات العنقودية للمصطلحات فضلاً عن مختلف خوارزميات التجريد . ولهذا العمل أهميته نظراً لأنه يختبر ويقارن بين مدى فعالية أنواع متعددة من سمات النظم في فهرس تجريبي على الخط المباشر.

تقييم الطرق الآلية :

من المهم بمكان ، عند النظر في تقييم الطرق الآلية، استكشاف ثلاثة جوانب، وهي كيف تم تقييم النظم الآلية؟ ثم ما هي النتائج التي أسفر عنها التقييم؟ وأخيراً ما كان لهذه النظم من أثر على صناعة الاسترجاع.

وعلى الرغم من اتساع مدى معايير تقييم النظم فإنه من الممكن التحقق من معالم الفئات العامة لعمليات التقييم. ومن المنطقي أن تكون الدراسات الرائدة في الطرق الآلية في الحقيقة هي دراسات الجدوى التي يتم إجراؤها لتحديد ما إذا كان من الممكن الحصول على مخرجات مناسبة اعتماداً على أسلوب معين في العمل. وعادة ما كانت النتائج يتم تسجيلها والإعلان عنها بواسطة الباحث الفعلي، كما كانت تشتمل على الانطباعات حول المخرجات، والصعوبات الحسابية والفكرية التي واجهت الحصول عليها . وكان هذا هو الحال فعلاً بالنسبة للدراسات الرائدة في التكشيف والاستخلاص بواسطة الآلات، والتي كانت تركز على الأساليب الإحصائية. كما كان هو الحال أيضاً بالنسبة للدراسات الخاصة بكل من الأساليب اللغوية وأساليب الذكاء الاصطناعي

المستخدمة في الكشف والاستخلاص وصياغة الاستفسارات بواسطة الآلات، والتي تميل لأن تكون تقارير عن الطرق التي تم تجريبيها والخبرة التي اكتسبها الباحث في تطبيقها. وخير مثال على ذلك عبارة سجلها مارشيتي Marchetti وبلكن (1991) Belkin : « لقد حاولنا في هذا البحث بيان كيف أفضى موقف نظري بعينه من قضية استرجاع المعلومات (وهو اعتبار التفاعل بؤرة عملية استرجاع المعلومات ، ونموذجاً للنصوص الفائقة لمراصد بيانات استرجاع المعلومات) أفضى إلى نتائج بحثية معينة، أفضت بدورها إلى وضع نظم عاملة فعلاً يَعد بتقديم مساعدات جوهرية في عمليات البحث التي يجريها المستفيد النهائي » . (P. 242) .

ومن الطرق الأخرى لتقييم النظم الآلية التجريبية مقارنة مخرجات هذه النظم بمواصفات معيارية معينة ، وعادة ما تكون هذه المواصفات المعيارية أدوات ناتجة يدوياً، ينظر إليها باعتبارها عالية الجودة . والواقع أن هذه هي طريقة التقييم الأساسية بالنسبة للاستخلاص الآلي. ومن ثم يقارن إدموندسون (1969) Edmundson مستخلصاته الناتجة آلياً « بالمستخلصات المستهدفة » وهي مقتطفات ناتجة عن التقاط المشاركين في التجربة ، يدوياً للجمال التي يرونها صالحة لأن يتضمنها المستخلص، وذلك بناء على توجيهات تحريرية. وهناك أيضاً حالات تتم فيها المقارنة بين المواصفات المعيارية الناتجة يدوياً ونتائج الكشف الآلي ونتائج بناء المكنز آلياً. وقد قارن إيفانز وآخرون (1991) Evans et al برنامجهم الخاص بالكشف الآلي بنتائج أداء المكشفين الذين يختارون يدوياً مصطلحات كشفية للوثائق نفسها، كما قارن بوركو وبرنك Borko (1963) and Bernick بين التوزيع البشري والتوزيع الآلي لمجموعة من الوثائق على مصطلحات لغة مقيدة صغيرة.

وهناك اتفاق عام في هذه الأنواع من عمليات التقييم على أن المخرجات الناتجة يدوياً تختلف جوهرياً عن تلك الناتجة بواسطة الآلات. ومن الممكن النظر إلى ذلك باعتباره نتيجة غير مشجعة، في حالة ما إذا كان الهدف من الطرق الآلية هو تقديم النتائج نفسها التي يمكن الحصول عليها بالطرق اليدوية. كما يمكن من ناحية أخرى النظر إلى الطرق الآلية باعتبارها مكملية للطرق اليدوية. ومن الممكن بوجه عام للطرق الآلية المعتمدة كما هو الحال فعلاً على المسوغ الأدبي، أن تعبر في الواقع عن جانب آخر للطابع التكاملي للغة الطبيعية واللغة المقيدة.

وأكثر السبل شيوعاً لتقييم التكشيف الآلي وبناء المكانز وصياغة الاستفسارات آلياً، هو التعرف على مستوى الأداء بناءً على الاستدعاء والتحقيق. وعادة ما يتم ذلك باستخدام مجموعة تجريبية، تتكون من عينة من الوثائق التي يغطيها أحد نظم الاسترجاع، بالإضافة إلى مجموعة من الاستفسارات وقرارات الصلاحية. ويتيح توافر هذه المجموعات التجريبية وتقاسمها، للباحثين القدرة على إجراء التجارب التي يمكن بها الحصول على القيم المقارنة للاستدعاء والتحقيق بالنسبة للنظم الآلية التي تعمل بالأساليب المراد دراستها أو بدون هذه الأساليب. ومن أكبر وأعرق المشروعات في هذا النوع نظام سمارت SMART الذي يتم فيه المقارنة بين مختلف الأساليب والمتغيرات، كتلك الخاصة بوزن المصطلحات، وأدوات التحكم في اللغة كالمكانز، وذلك بقصد التعرف على ما لهذه الأساليب من أثر في كفاءة الاسترجاع (Salton, 1968).

وتلخص سبارك — جونز (Sparck-Jones (1981 عدداً كبيراً من الجهود المتصلة بالدراسات التجريبية في استرجاع المعلومات، وتخلص إلى أن استخدام الكلمات المجردة المعتمدة على اللغة الطبيعية، يكاد يتساوى في كفاءته مع غيره من الأساليب اللغوية الأكثر تعقيداً كالروابط والأدوار. هذا بالإضافة إلى أنها قد انتهت إلى أن الطرق الآلية لم تكن أكثر فعالية من الطرق اليدوية. إلا أننا ينبغي أن نراعي بعض الحذر في تفسير هذه النتائج، وذلك للأسباب التالية:

١ - إن كثيراً من تجارب الاسترجاع يتم إجراؤها على مجموعات صغيرة (تتراوح ما بين بضع مئات وعدة آلاف من الوثائق). وهناك ما يدل على أنه من الممكن لمتغير الحجم أن يكون له أثر لا يستهان به على مقاييس الأداء (Blair and Maron, 1985).

٢ - كذلك أجري الكثير من التجارب اعتماداً على أعداد قليلة من الاستفسارات، بلغت في بعض الأحيان عشرة استفسارات أو أقل. ومن الصعب أيضاً معرفة ما إذا كان من الممكن للنتائج المستخلصة في ظل هذه الظروف، أن تتخذ أساساً للتنبؤ بأداء النظم العاملة التي تتعامل مع أعداد كبيرة من الاستفسارات المتنوعة.

٣ - هناك تفاوت كبير في النتائج؛ حيث يمكن، في بعض الأحيان، لأرقام الاستدعاء والتحقيق أن تكون مختلفة تمام الاختلاف، بين كل من مستولي البحث المتعاملين مع المجموعة نفسها، وبالنسبة لمستول البحث نفسه الذي يستخدم طريقة تجريبية معينة في عدة مجموعات مختلفة.

وتختتم سبارك - جونز (Sparck-Jones (1981) دراستها الوصفية التحليلية لعشرين عاماً من تجارب استرجاع المعلومات، قائلة: إنه من الصعب حتى الآن التنبؤ بمدى الكفاءة بدقة ، نظراً لتعقد نظم الاسترجاع. وربما كان من الممكن لهذه الحقيقة مضافاً إليها النتيجة العامة التي تؤكد أن معظم الطرق التي تم تجربتها لم تثبت تفوقاً يذكر على الطرق اليدوية من حيث الاستدعاء والتحقيق ، ربما كان من الممكن لذلك أن يفسر لماذا لم يسلك سوى قليل جداً مما حدث في تجارب استرجاع المعلومات سبيله إلى بيئة الممارسة الفعلية؟.

وتتناول المناقشات الحديثة لقضية التقييم في استرجاع المعلومات الكثير من المشكلات التي تعرضنا لها في هذا القسم، ومن بينها حجم المجموعة (Salton, 1992) والصعوبات التي تكتنف الوصول بالإجراءات التجريبية إلى النظم العاملة الضخمة (Ledwith, 1992) . وتدل الجهود التجريبية والتنفيذية الجارية على أن هذه القضايا تعالج الآن على نحو جاد:

١ - في إطار مبادرتي تِرِك TREC وتبستر TIPSTER تجري دراسة أنواع مختلفة من الطرق التجريبية ، اعتماداً على مرصد بيانات بالغ الضخامة (٢ جيجا بايت) (Harmon, 1992).

٢ - سلكت إحدى الطرق التجريبية ، وهي مخرجات الاسترجاع المرتبة طبقياً والناجمة عن إجراءات الوزن الإحصائي ، سبيلها في عدد قليل من المنتجات التجارية (Visschedijk and Gibb, 1993).

الفصل الثاني عشر

الاتجاهات والتطورات المستقبلية المحتملة

من الممكن أن نجد في الشكل رقم (٨٠) بعض الاتجاهات المهمة التي تؤثر في نظم استرجاع المعلومات ، حيث يصور هذا الشكل تقسيماً ثنائياً بين التقنيات (الأدوات أو الوسائل أو الوسائط الخاصة باختزان المعلومات وتنظيمها وبثها) والبيئة التي تستخدم فيها التقنيات (المحتوى الفكري لمصادر المعلومات، والمنظمات والأفراد الذين يفيدون من هذه المصادر). ويمكن للخص عام للنقاط التي يعبر عنها الشكل أن يهييء المجال لما يلي من مناقشات متعمقة .

لقد كان هناك نمو مطرد في منتجات وأشكال تقنيات المعلومات ، بما في ذلك الحاسبات الآلية ، ومراصد البيانات ، والبرمجيات ، ووسائط الاختزان (التي كانت تقتصر من قبل على المطبوعات والمصغرات الفيلمية ، ولكنها أصبحت الآن تشمل الأشكال القابلة للقراءة بواسطة الآلات ، كالأسطوانات المغنطة والأسطوانات الضوئية). كما كان هناك أيضاً اتجاه واضح نحو المشابكة ، حيث تتاح لكثير من المنظمات والأفراد فرصة التعامل مع أنواع متعددة من المصادر النائية عن طريق شبكات الاتصالات بعيدة المدى مثل البنتنت Bitnet والإنترنت Internet .

وكما هو الحال بالنسبة لمعظم الابتكارات المهمة ، فقد استخدمت الوسائل والأشكال الإلكترونية استجابة للاهتمامات وتلبية للاحتياجات وحلاً للمشكلات . إلا أنه نظراً لما لها من أثر على من يفيد منها من المنظمات والأفراد غالباً ما تستخدم بطرق لم تكن مقصودة على الإطلاق؛ فقد تتبع شوركين (1984) Shurkin على سبيل المثال تطور

الحاسب الآلي ، حيث يسجل أنه تطور لكي يؤكد جدواه كأداة ، في الوقت نفسه الذي يعمل فيه على زيادة سرعة إجراء الأعداد الضخمة من العمليات الحسابية التي كان المهندسون بحاجة إلى إجرائها في غضون الحرب العالمية الثانية. وقد أصبح منذ ذلك الوقت جزءاً لا غنى عنه في جميع ضروب الأنشطة والأعمال والمؤسسات تقريباً (بما في ذلك المكتبات) كما يستخدم الآن في جهود متنوعة كمراجعة الأرصدة في مجال إدارة الأعمال ، وفي بحث نصوص الدساتير والقوانين واسترجاعها، وفي تنظيم سجلات المستشفيات، وفي إدارة الشئون المالية (Stallings, Hutchinson, and Sawyer, 1988) . ويتعامل الأفراد الآن على اختلاف فئاتهم وتفاوت مستوياتهم وقدراتهم مع سلسلة عريضة من المصادر القابلة للقراءة بواسطة الحاسبات الآلية، في كثير من الأماكن المختلفة. وتثير هذه الظاهرة عدداً كبيراً من القضايا الجديرة بالاهتمام، بما في ذلك أخلاقيات استخدام الحاسبات في مهام معينة (كمراقبة أداء العاملين مثلاً أو تتبع تطور أرصدة عملاء المصارف) (Marx and Sherizen, 1989) كما تثير أيضاً تساؤلات مهمة حول من يحق له ومن لا يحق له التعامل مع البنية الأساسية للمعلومات، وقد جعلت كل هذه القضايا من أهداف استرجاع المعلومات الرامية إلى الوصول إلى المعلومات التي تدعو الحاجة إليها، واقتطاف هذه المعلومات من المصادر الملائمة، مهمة متزايدة الصعوبة.

| تقنيات المعلومات | بيئات المعلومات |
|---|---|
| التحول من الأشكال الورقية إلى الأشكال الإلكترونية للوثائق . | قضايا الحفظ والصيانة وحقوق التأليف والنشر . |
| تزايد العتاد والبرمجيات ، من مراصد البيانات، والبرامج والنظم، ووسائط المطبوعات. | سوق المعلومات . |
| المشاركة ، وتوزيع جهد الحاسبات ، وتصوير الوثائق عن بعد . | قضايا التعامل مع مصادر البيانات والمكتبة الافتراضية . |

الشكل رقم (٨٠) : مصادر وبيئات تقنيات المعلومات - الاتجاهات والقضايا .

نمو مصادر المعلومات وتقنيات المعلومات :

من الممكن تعريف تقنيات المعلومات بأوسع معانيها بحيث تشمل ما يلي :

- ١ - العتاد (أو الوسائط) الخاص باختزان المعلومات ، بما في ذلك المطبوعات، والمصغرات الفيلمية، ومختلف الوسائل الإلكترونية، كالأسطوانات المغنطة والأسطوانات الضوئية.
 - ٢ - العتاد الخاص بتجهيز المعلومات ونقلها ، بما في ذلك آلات التصوير الضوئي ، وأجهزة تصوير الوثائق عن بعد ، والحاسبات، وأجهزة ربط الحاسبات بشبكات الاتصالات.
 - ٣ - الوحدات الفكرية والمادية التي تقدم بها المعلومات، وهذه تشمل المطبوعات الأولية التي تضم النصوص الكاملة، كالدوريات والكتب والتقارير، فضلاً عن المصادر الثانوية ومصادر الدرجة الثالثة، كالموسوعات ، والكشافات الموضوعية ، وفهارس المكتبات التي تعرف بأوعية الإنتاج الفكري الأولي. وغالباً ما تتخذ المطبوعات شكل الكتب بينما تضم الأوعية الإلكترونية ملفات تعرف بمراصد البيانات.
 - ٤ - أساليب وطرق استرجاع وعرض المواد من المصادر الأولية والمصادر الثانوية. ويمكن بالنسبة للمطبوعات استخدام الفهارس المطبوعة أو الفهارس البطاقية في البحث اليدوي. أما بالنسبة للحاسبات الآلية فإن هناك برمجيات الواجهات التي تشتمل على الأوامر والتعليمات التي يتوسل بها المستفيد لاسترجاع المعلومات وعرضها.
- وقد حدث في غضون العقود الأربعة الأخيرة تزايد كبير في مصادر المعلومات وتقنيات المعلومات، تمثل في :
- ١ - نمو عدد المطبوعات وخاصة الدوريات المتخصصة وما تشتمل عليه من مقالات.
 - ٢ - نمو عدد المصادر الثانوية (كالكشافات والفهارس مثلاً) بمختلف أشكال الوسائط (كالمطبوعات والأسطوانات الضوئية مثلاً) وكذلك نمو عدد المداخل الواردة في هذه المصادر.

٣ - نمو عدد الحاسبات الآلية التي تستخدم في أنواع شتى من المؤسسات ومن جانب الأفراد.

٤ - النمو في أعداد ونوعيات منتجات البرمجيات المتاحة على المستوى العام، وكمنتجات تجارية ونظم.

وتتسم هذه الظواهر جميعاً بالنمو الآسي . وكما ذهب ميدوز وآخرون Meadows et al. (1972) فإن « الكم يمثل نمواً أسياً عندما يزداد بنسبة مئوية ثابتة من الكل في فترة زمنية ثابتة » (P. 27) . ويعني ذلك أن النمو يتأثر إلى حد بعيد برقم الأساس ؛ فإذا كان هذا الرقم صغيراً فإن الكمية المضافة يمكن أن تكون صغيرة ، أما عندما يرتفع الأساس إلى أرقام أكبر فإن الكمية المضافة تتزايد بشكل أكبر.

ومن أشهر محاولات رصد النمو الآسي ما سجله درك دي سولا برايس Derek de Solla Price ، حيث يرصد كتاباه (1963) *Little Science, Big Science* و (1961) *Science Since Babylon* نمط النمو الآسي في النشاط العلمي ، وخاصة في الدوريات العلمية وما تشتمل عليه من مقالات أولاً، ثم في مصادر التكشيف والاستخلاص التي ظهرت لإحاطة الباحثين بأحدث التطورات الجارية في مختلف المجالات. وقد بين دي سولا برايس أن نمو الإنتاج الفكري العلمي لم يكن أسياً فحسب ، وإنما كان سريعاً وبشكل مذهش أيضاً. وذهب فعلاً إلى القول بأن الإنتاج الفكري بالنسبة لمعظم مجالات النشاط العلمي يتضاعف كل عشر سنوات (Price, 1963, P. 6) .

ولا يقتصر التزايد الآسي على عدد المطبوعات الأولية والثانوية ، وإنما يشمل أيضاً عدد النظم التي يتعين على المرء أن يعرف كيف يفيد منها للوصول إلى كل هذه المطبوعات. ولا يقل عن ذلك ضخامة تزايد أعداد متعهدي الخط المباشر. وبوابات العبور، ومراصد البيانات ، وإن كان هناك ما يدل على أنه قد لا يكون بالسرعة التي كان عليها في مطلع ثمانينيات القرن العشرين (Williams, 1990) .

كذلك شهدت الفئات العامة لعتاد الحاسبات الآلية وبرمجياتها نمواً أسياً. ولا يصدق ذلك على عدد الحاسبات المشتركة فحسب، وإنما على مختلف الخصائص المهمة لعتاد الحاسبات وبرمجياتها أيضاً. وبعبارة أخرى ، فإن سرعة الحاسبات، ومدى

تعقدتها، وسعتها الاختزانية قد تزايدت بمعدلات أسية. إلا أنه على الرغم من الارتفاع المستمر في هذه الأرقام ، هناك أرقام أخرى تتناقص وبالمعدل السريع نفسه، وهذه تشمل تكلفة وحدات العتاد وأحجام هذه الوحدات. ومن ثم فإنه من الممكن للمرء ، بمبلغ مالي معين شراء حاسب آلي شخصي يتمتع بقدرات أكبر مما كان عليه منذ بضع سنين ، لا من حيث السرعة في تنفيذ التعليمات فحسب ، وإنما في مدى تعقد البرامج التي يستطيع تنفيذها وكميات البيانات التي يمكن اختزانها أيضاً .

| العقد | الدفعات | تقاسم الوقت | المكتسبي | الشبكات |
|---------------|------------------|--------------------|---------------|-----------------|
| الموقع | السنتينيات | السبعينيات | الثمانينيات | التسعينيات |
| المستخدمون | غرفة الحاسب | غرفة المنفذ | مسطح المكتب | متحركة / متنقلة |
| حالة المستخدم | الخبراء | الاختصاصيون | الأفراد | المجموعات |
| البيانات | تابع | الاعتماد على الغير | الاستقلال | الحرية |
| الهدف | هجائية / رقمية | نصوص ، كمية | أبناط ورسوم | النص والصوت |
| نشاط المستخدم | الحساب | الوصول | التقديم | الاتصال |
| همزة الوصل | التثقيب | التذكر والرقن | يرى ويشير | يسأل ويخبر |
| التطبيقات | (التقديم) | (التفاعل) | (يشغل) | (يفوض) |
| | المعدات المساعدة | المنافذ | أجهزة المكاتب | شبكات الاتصالات |
| | محلية | معيارية | عامة | متنوعة |

الشكل رقم (٨١) : الأطر الأربعة لاستخدام الحاسبات .

وقد افترضنا في هذا الكتاب وجود نموذج لاستخدام الحاسب الآلي، يتكون من نظام لتقاسم الوقت ، يرتبط فيه المستخدمون ويتصلون بمراصد البيانات المحملة على حاسب عملاق، أو يتكون من نظام مبني على مجموعة من الحاسبات متناهية الصغر أو مجموعة من محطات العمل القوية، يتعامل فيه المستخدمون مع البيانات محلياً، على

أسطوانات ضوئية مكتنزة مثلاً. ويوضح الشكل رقم (٨١) هذه الأطر الخاصة باستخدام الحاسبات ويقدم سياقاً عاماً للتعرف على ما كان عليه استخدام الحاسبات وفي أي اتجاه يسير. ويصف تسلر (1991) Tesler تتابع الأحداث الكامنة وراء الجهود التي شهدتها كل عقد من العقود على النحو التالي:

« لقد اخترع الإطار الأصلي للحاسب في نهاية أربعينيات القرن العشرين، حين صممت الآلة الحاسبة القابلة للبرمجة كأداة في متناول المهندسين ، ثم أصبحت قابلة للاستخدام على المستوى التجاري في خمسينيات القرن نفسه. ثم حدث أول تحول في الستينيات ، حين اتخذ الحاسب الآلي كأداة لتجهيز البيانات من جانب المؤسسات. ثم جاء التحول الثاني في السبعينيات عندما بدأ تقاسم خدمات الحاسب بين عدد كبير من المشتركين. وقد حدثت الطفرة الثالثة في الثمانينيات حيث حولت الحاسب إلى أداة للارتفاع بإنتاجية الأفراد ، تتوافر بالمكاتب. أما التحول الرابع فيحدث الآن ، ومن بشائره تزايد المشابكة بين أجهزة الحاسبات الشخصية ، وتقويم الجيب الإلكترونية، والأجهزة المتنقلة التي أسميها بأشبه الحاسبات Pericomputers . ولأشبه الحاسبات هذه أهميتها المرتبطة بالوظائف المحددة التي يمكنها تنفيذها بمفردها ، فضلاً عن قدرتها على الاتصال برصيد المعلومات المسجل على وسائط إلكترونية » . (PP. 87,88) .

ومن الممكن الاتصال بعالم المعلومات هذا عن طريق منظومة هائلة قوية من شبكات الاتصالات بعيدة المدى . ولهذه البيئة المتشابكة أهميتها القصوى لسببين :

١ - أنها تكفل إمكانية الاتصال بأنواع المصادر شتى من جانب المستفيدين الأفراد.

٢ - أنها تكفل للمستفيدين كمجموعات القدرة على الاتصال ببعضهم

البعض (Quarterman, 1990) .

ويتم الاتصال بالمصادر عن طريق شبكات الاتصالات بعيدة المدى مثل البتنت Bitnet والإنترنت Internet . والبتنت شبكة تعاونية ، بلغ عدد المستفيدين منها عام ١٩٨٨ م ، ٢٣٠٠ مضيف في اثنتين وثلاثين دولة . أما الإنترنت ، التي تعد فعلاً شبكة للشبكات فتخدم نطاقاً جغرافياً شاسعاً يشمل كلاً من الولايات المتحدة الأمريكية وكندا وأوروبا وآسيا (*) . وقد تزايدت الاستفادة من الإنترنت بسرعة مع مرور الوقت ، من ٢١٣ مضيفاً عام ١٩٨١ م إلى ٧٢٧٠٠٠ مضيف في مطلع عام ١٩٩٢ م (Lottor, 1992) .

(*) وصلت أيضاً أفريقيا والاقويانوس . (المترجم) .

أما المصادر التي يمكن الاتصال بها عبر هذه الشبكات فلا تقتصر على مراصد البيانات والبرمجيات اللازمة للإفادة منها ، وإنما تشمل أيضاً موارد العتاد كحيز الاختزان والطابعات . وعلى ذلك ، فإن بيئة استخدام الحاسبات في تسعينيات القرن العشرين لا مركزية إلى حد بعيد ، إلا أنها مترابطة فيما بينها ، على عكس مركزية الستينيات والسبعينيات ، والاستقلال النسبي الذي كفلته الحاسبات المكتبية متزايدة القوة في الثمانينيات . ولقد أصبح بإمكان المستفيد في التسعينيات التعامل مع موارد بكفاءة عالية ، فضلاً عن دعم هذه الموارد بمنظومة ضخمة من موارد العتاد والبرمجيات التي يمكنه الاتصال بها والتعامل معها عبر الشبكات.

ومن العوامل المهمة الأخرى في بيئة التسعينيات المترابطة فيما بينها قدرة الأفراد على الاتصال إلكترونياً ببعضهم البعض . وهناك أنواع متعددة من الخدمات المتاحة لهذا الغرض ، ومن بينها البريد الإلكتروني ونظم الائتثار بواسطة الحاسبات والتي تتيح فرصة الحوار والنقاش بين عدد كبير من المشاركين . ويرتبط بهذه الإمكانية تزايد احتمالات تضافر جهود الجماعات إلكترونياً ، في بعض المشروعات . ولب هذه الفكرة هي « الوثيقة المشتركة » التي يمكن لكل فرد أن يتعامل معها بالإضافة أو التعديل أو الحذف (Schrage, 1990) . ويمكن النظر إلى هذه الوثيقة في بعض الأحيان كوثيقة مركبة ، تتكون من أشكال متنوعة من الوسائط ، كالنصوص والصور ، والرسومات ، والجداول ، والإشارات السمعية والبصرية (Heller, 1991) . وتدل الوثائق المركبة أيضاً على جانب مهم آخر من جوانب بيئة استخدام الحاسبات في التسعينيات وما بعدها: فبينما كانت مراصد بيانات الستينيات والسبعينيات نصية ورقمية في الأساس ، فإن مراصد البيانات الحديثة « تفاعلية » وتشتمل على الصور والأصوات والوسائط الأخرى (Lucky, 1989; Cawkell, 1992) .

استرجاع المعلومات وأطر استخدام الحاسبات :

لا عجب أن يساير تطور استرجاع المعلومات أطر استخدام الحاسبات التي عرضنا لها بالوصف في القسم السابق (وكذلك في الفصل الثاني) . فقد كان هناك عدد من نظم التجهيز على دفعات في ستينيات القرن العشرين ، ومن بينها خدمة المدلرز MEDLARS بالمكتبة القومية للطب ، ونظام مركز الحاسب الآلي للمكتبات على الخط المباشر Online (OCLC) Computer Library Center الذي كان يقدم بطاقات الفهرسة للمكتبات

اعتماداً على رصيده من تسجيلات الفهرسة القابلة للقراءة بواسطة الآلات ، في صيغ مارك MARC . وقد أسفرت التطورات التي حدثت في تقنيات الاختزان والاسترجاع والاتصالات بعيدة المدى في السبعينيات ، عن تطور هذه الخدمات ، وعن إدخال نظم جديدة تعتمد على الاسترجاع التفاعلي على الخط المباشر. فقد تحول المدلرز حينئذ إلى المدلاين MEDLINE (المدلرز على الخط المباشر) كما أصبحت خدمة مركز الحاسب الآلي للمكتبات OCLC التي كانت تعمل على دفعات، متاحة على أساس تقاسم الوقت، حيث أصبح بإمكان المهرسين إدخال التسجيلات وتحريرها ومراجعتها واسترجاعها بطريقة تفاعلية. هذا ، بالإضافة إلى إدخال كثير من نظم الاسترجاع الأخرى ، كما بدأت الفهارس المتاحة على الخط المباشر تظهر في المكتبات الأكاديمية الكبرى، التي كانت لديها القدرة على تنفيذ هذه الفهارس وصيانتها اعتماداً على الحاسبات العملاقة .

وقد ظل التعامل مع مراصد البيانات الضخمة عن طريق عدد من خدمات الاسترجاع الوراقية ، إحدى الدعائم الأساسية في استرجاع المعلومات خلال السبعينيات. إلا أنه مع تناقص أحجام الحاسبات وانخفاض تكلفتها ، أصبح بإمكان كثير من المكتبات بما فيها المكتبات الصغيرة ، إنشاء وصيانة الفهارس المتاحة على الخط المباشر، اعتماداً على النظم التجارية الجاهزة المتاحة ، وهي حزم العتاد والبرمجيات التي تم تصميمها وتسويقها لمختلف قطاعات العمل بالمكتبات ، كتسجيل الإعارات ، وتبادل الإعارة بين المكتبات ، والبحث في الفهارس . وقد أسفر المزيد من التصغير في مكونات الحاسبات، عن ظهور الحاسبات متناهية الصغر ، التي أدت إلى جعل تكلفة استخدام الحاسبات في متناول قطاعات عريضة من المكتبات . وقد أصبح من الممكن الآن وبوجه عام للمكتبات ، إنشاء وصيانة العديد من مراصد البيانات المحلية ، المكونة من التسجيلات التي يتم الحصول عليها من مصادر أخرى ، أو تلك التي يتم تجهيزها محلياً، والتي تكفل استرجاع جزء من مقتنيات المكتبة ، أو جميع المقتنيات (Tenopir and Lundeen, 1988) . هذا بالإضافة إلى أن توافر مراصد البيانات التجارية على أسطوانات

ضوئية مكتنزة ، قد جعل المقتنيات الضخمة ، ومن بينها تلك التي كان يتم التعامل معها على الخط المباشر عن طريق الحاسبات العملاقة ، في متناول المكتبات. وكما هو الحال بالنسبة لمراصد البيانات والوسائط الأخرى ، فإن نمو أعداد الأسطوانات الضوئية المكتنزة التي تباع للمكتبات وتستخدم في هذه المكتبات ، في ارتفاع مستمر . فقد تبين لكل من نيكولز وفان دن إلشاوت Nicholls and Van Den Elshout عام ١٩٩٠م، أنه كان

هناك حوالي ٥٠٠ مرصد للبيانات في متناول المكتبات على أسطوانات ضوئية مكتنزة ، وفي عام ١٩٩٣ م ارتفع الرقم إلى أكثر من ١٢٠٠ .

ومما هو جدير بالاهتمام في هذا الصدد ملاحظة كيف غيرت التقنيات من علاقة المكتبات ببيئتها ؛ فقد شاركت المكتبات في الجهود التعاونية متأثرة بالأتمتة في المقام الأول. فقد كانت المكتبات قبل إدخال الحاسبات الآلية ، تعمل كوحدات مستقلة نسبياً ، على الرغم من أنه كانت هناك أمثلة قليلة للجهود التعاونية الضخمة ، مثل الـ *National Union Catalog* ، الذي نشأ وحظي بالصيانة لأغراض تقاسم الموارد. ومع الاتجاه نحو استخدام الحاسبات والاسترجاع على الخط المباشر، تزايد ارتباط المكتبات بالموارد المركزية النائية ، كتلك التي توفرها خدمات الاسترجاع الوراقى والمرافق الوراقية . ونظراً لتناقص تكلفة الحاسبات وما صاحبه من تزايد قدرتها التجهيزية وطاقتها الاختزانية ، أصبح من الممكن لأعداد متزايدة من المكتبات تقديم خدمات مستقلة نسبياً اعتماداً على مرصد البيانات التي نشأت محلياً وتلك التي يتم الحصول على ترخيص التعامل معها من المتعهدين . إلا أن أحدث التطورات التي تحققت فعلاً هو تحول الكثير من المكتبات ومراكز المعلومات إلى أطراف في بيئة استخدام الحاسبات في العقد الأخير من القرن العشرين ، وهي بيئة لا مركزية ، موزعة ولكنها مترابطة فيما بينها. ويسجل دليل بارون (Barron, 1992) *The Barron's Guide* على سبيل المثال أكثر من ٣٠٠ فهرس متاح على الخط المباشر ، عن طريق الإنترنت ، ويمكن الوصول إليها والبحث فيها من جانب كل من تتوافر له إمكانات التعامل مع إحدى الشبكات . هذا بالإضافة إلى أنه قد أصبح من الممكن الاتصال بأعداد متزايدة من مصادر المعلومات المتنوعة عن طريق الشبكة، من خلال وسطاء أو خدمات مثل وائز (WAIS (Wide Area Information Server وجوفر GOPHER والشبكة الدولية WWW (World Wide Web) (Krol, 1992) (*)

التحول من المطبوعات إلى الوثائق الإلكترونية :

قدم لانكستر (Lancaster (1978 منذ ما يزيد على العقد رؤيته للمجتمع اللاورقي، كما سجل توقعاته لما يمكن أن يكون له من أثر على المكتبات (Lancaster, 1982) . كذلك توقع احتمال وقوع أحداث معينة في عدد من الجوانب لخصها على النحو التالي (Lancaster, 1985, P.554)

(*) لمزيد من المعلومات في هذا الموضوع راجع :

هشمت قاسم. الإنترنت ومستقبل خدمات المعلومات. دراسات عربية في المكتبات وعلم المعلومات، ع ٢، ١٩٩٦ م.
ص ص ٤٤ - ٨٨ (المترجم).

الجانب الأول : نوعيات المطبوعات :

تأثر مختلف الأنواع بالإلكترونيات في تسلسل منطقي ، يعكس الاحتياجات النسبية والمزايا المتوقعة .

الجانب الثاني : استخدام الحاسبات في النشر (في تتابع تطوري)

- أ - الطباعة على الورق .
- ب - لإتاحة البديل الإلكتروني للمطبوعات الورقية .
- ج - لإنتاج أوعية جديدة ليس لها نظير مطبوع .
- د - لإدخال أشكال من الأوعية بإمكانات جديدة تماماً .
- هـ - للعمل على استبدال الأشكال المطبوعة الحالية .

الجانب الثالث : مقدار البيانات المختزنة :

- أ - الإشارات الوراقية فقط .
- ب - المستخلصات .
- ج - النصوص الكاملة .

الجانب الرابع : نوعية خدمات المعلومات المتأثرة :

- أ - البحث في الإنتاج الفكري .
- ب - الرد على الاستفسارات .
- ج - الإمداد بالوثائق .

الجانب الخامس : المجال الموضوعي :

- أ - العلوم والتقانة .
- ب - العلوم الاجتماعية .
- ج - الإنسانيات والأدب .

الجانب السادس : نوعية المتلقي المتأثر :

- أ - المؤسسات المستفيدة .
- ب - الأفراد .

أو وفقاً للمكتبات :

- أ - المكتبات المتخصصة .
- ب - المكتبات الأكاديمية .
- ج - المكتبات العامة .

وكثير مما أشرنا إليه هنا إما أن يكون قد تحقق فعلاً أو أن يكون في سبيله للتحقق، وهناك كثير من الشواهد على أن الاتجاهات المتوقعة سوف تستمر . ومن ثم فإنه من المفيد التعرف على الموقف الراهن للمجال واحتمالات تطوره في المستقبل، وتصور ما يمكن أن يكون لهذه الاتجاهات من أثر على استرجاع المعلومات وبثها .

لقد قدم الفصل الثاني تطور نظم الاسترجاع على الخط المباشر ، باعتباره قد بدأ كنتاج جانبي لعملية النشر ، أي التنضيد بواسطة الحاسب الآلي ، الذي حل محل تقنيات أخرى أعلى تكلفة وأكثر استنزافاً للوقت ، وقد أسفرت تلك العملية عن شكل قابل للقراءة بواسطة الآلات لمراصد البيانات التي كانت تستخدم لاسترجاع المعلومات على دفعات ثم على الخط المباشر . وقد ظهرت الآن الأوعية التي لا تتوافر إلا في شكل قابل للقراءة بواسطة الآلات . ويمكن في بعض الحالات أن يكون مرصد البيانات مغرقاً في التخصص بحيث لا يمكن أن يحقق عائداً في سوق المطبوعات . ويمكن في حالات أخرى أن تكون المادة مؤقتة إلى حد بعيد (كما هو الحال مثلاً بالنسبة لمراصد البيانات الخاصة بالنصوص الكاملة للنشرات الإخبارية) أو تكون المادة سريعة التغير (كما هو الحال مثلاً بالنسبة لبعض بنوك المعلومات الرقمية ، كتلك الخاصة بالبورصة وسوق المال) ، أو تكون البيانات معبرة عن التطورات التي استجدت والتي ينبغي أن تصل إلى المستفيدين بسرعة (كالدورية *The Online Journal of Current Clinical Trials* التي تصدر عن مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC والجمعية الأمريكية لتقدم العلوم) (Rogers, 1991) . أضف إلى ذلك أنه يمكن للأوعية التي تشتمل على الإشارات الوراقية فقط ، أو الإشارات الوراقية فضلاً عن المستخلصات ، والنصوص الكاملة ، أن تتوافر في الشكل المطبوع فقط ، أو في الشكل المطبوع وعلى الخط المباشر ، أو على الخط المباشر فقط . هذا بالإضافة إلى أن هذه الأوعية تغطي جميع المجالات ، من العلوم والتقانة إلى الأدب المبسط (Williams, 1985) .

وقد اقترحت أشكال النشر ذات الإمكانات الجديدة غير المسبوقة ، في وقت مبكر من تاريخ استرجاع المعلومات ، وقد أسهمت هذه الأشكال منذ ذلك الوقت ، وبشكل ملحوظ في بحوث استرجاع المعلومات وتطبيقاته . وربما كان فانيفار بوش Vannevar Bush أول من أدرك أهمية الوسائل الجديدة بالنسبة للتنظيم الفعال للمعلومات واسترجاع المعلومات؛ ففي مقالة نشرت في مجلة *Atlantic Monthly* تصور بوش أداة أسماها «الممكس memex» يمكن للمستفيد أن يختزن بها ملفاً خاصاً من الوثائق (Bush, 1945). وكان من الممكن للممكس أن تتمتع بقدرات استرجاع قوية مرنة ، يمكن أن تكفل للمستفيد القدرة على بناء منظومة من الارتباطات بين المواد ذات الأهمية ، وعلى تسجيل تعليقاته على هذه المواد.

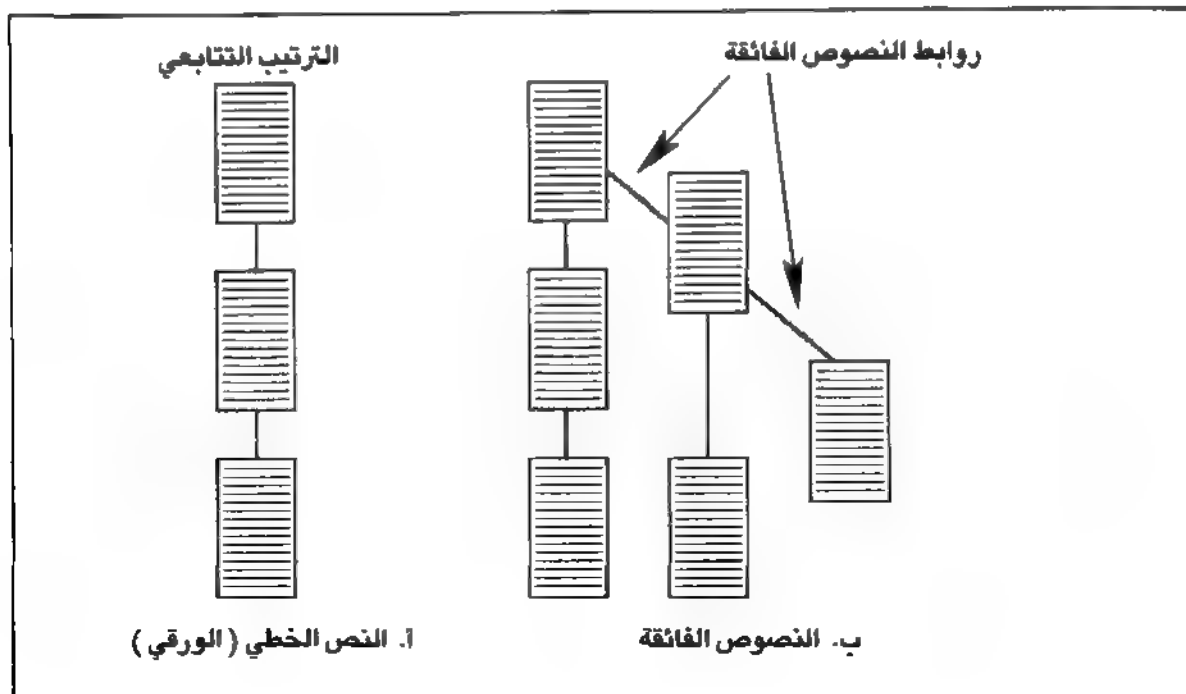
وبناء على هذه المقالة ارتبط اسم بوش بكل من بدايات الأساليب الحديثة لاسترجاع المعلومات ، ونشأة فكرة النصوص الفائقة التي اتسعت وتطورت على يدي نلسون (1967) Nelson فيما بعد . فالنصوص الفائقة تختلف عن الوثائق التقليدية ، سواء في شكلها المطبوع أو في الشكل القابل للقراءة بواسطة الآلات ، وذلك من عدة أوجه مهمة ؛ فالنص التقليدي خطي وثابت ، حيث تستقر الكلمات في الوثائق بطريقة تتابعية . وتظل هذه الكلمات كما هي بمجرد كتابتها ونشرها . وأخيراً ، فإنه على الرغم من إمكان الوصول إلى الوثائق التقليدية بعدة طرق ، من بينها البحث بالكلمات أو أجزاء الكلمات ، واسترجاع نصوصها كاملة أو أجزاء من هذه النصوص استجابة لاستفسار ما ، فإن كل نص أنتجه مؤلف ، إنما هو في الواقع له سياق زمني المحدد ، كما أنه مستقل بذاته . ويمكن لهذا النص أن يسترجع أو لا يسترجع مع غيره من الوثائق التي تشتمل على مصطلحات تضاهي الاستفسار .

أما في النصوص الفائقة فإن العناصر النصية ترتبط فعلاً ببعضها البعض بأكثر من طريقة . ومن ثم فإن النظام لا يتوافق به نصوص قائمة بذاتها فحسب ، وإنما يربط بين هذه النصوص وبعضها البعض ، بحيث يمكن للمستفيد أن ينفذ إلى ما يريده منها . ويسمى نيلسن (1990) Nielsen النصوص الفائقة « بالكتابة غير التتابعية » ، ويصفها على النحو التالي :

« ... نص موجه ، حيث تشتمل كل نقطة ارتكازية على قدر معين من النص أو أية معلومات أخرى . وترتبط النقاط الارتكازية ببعضها البعض بروابط موجهة . ويمكن للنقطة الارتكازية ، في معظم نظم النصوص الفائقة أن تخرج منها عدة روابط ثم تقترن كل منها بأحد العناصر الصغيرة في النقطة الارتكازية يسمى المرساة أو المركز (الهلب) Anchor . وعندما يتعامل المستفيدون مع أحد هذه المركبات فإنهم يتتبعون الرابطة المقترنة به إلى النقطة الارتكازية التي تنتهي عندها ، وبذلك يقومون بالملاحة في شبكة النصوص الفائقة . ويعود المستفيدون على نفس الدرب باتباع الروابط التي استخدموها في الملاحة في الاتجاه العكسي . ومعالم الحدود Landmarks نقاط ارتكازية بارزة على نحو خاص في الشبكة ، كأن تكون على سبيل المثال من الممكن الوصول إليها مباشرة من أي نقطة ارتكازية (أو من جميع النقاط الارتكازية) » . (P. 298) .

ويوضح الشكل رقم (٨٢) البنية النظرية للعناصر كما تتصل ببعضها البعض في النظم الورقية الخطية ونظم النصوص الفائقة . ولأغراض هذا المثال فإنه لا يهم ما إذا كانت الوحدات المفردة في كل حالة وثائق كاملة أو أجزاء من وثائق ، أو إشارات ورقية ،

أو صوراً أو أصواتاً . وما يهم فعلاً أن ترد العناصر في التتابع الخطي تلو بعضها في تسلسل . وربما كان من الممكن في مثل هذه النظم ، بالنسبة للمستفيد أن يغير ترتيب العرض . فإذا كانت العناصر على سبيل المثال نصوصاً مصحوبة بتاريخ النشر ، فإن ترتيبها يمكن أن يكون هجائياً أو زمنياً ، إلا أنها رغم ذلك يمكن أن تقدم تتابعياً .



الشكل رقم (٨٢) : مقارنة بين الوثائق التقليدية ووثائق النصوص الفائقة .

وفي نظام النصوص الفائقة ، ترتبط العناصر المفردة (النقاط الارتكازية) بطرق كثيرة بالروابط . ويؤدي ذلك إلى وجود « وثيقة فائقة » يتم تصميمها جزئياً في الواقع بواسطة مصمم النظام الذي أنشأ الروابط ووضعها ، وجزئياً بواسطة القارئ ، الذي ينفذ إلى تلك الروابط على النحو الذي يريده عند استخدام النظام ، ونظراً لأن المسارات التي يتم اتخاذها في كل حالة هي التي تقرر بنية الوثيقة الفائقة ، فإنه لا يمكن الحصول على نسخة ورقية من الوثيقة الفائقة ، حيث تختلف في كل مرة يتم فيها استخدام النظام .

ولنظم النصوص الفائقة والوسائط الفائقة (أي النظم التي تشتمل على الأصوات والصور فضلاً عن النصوص) مستقبل واعد . فقد كانت هذه النظم موضوعاً لكثير من مشروعات البحث والتطبيق التي نشرت تقارير عنها في الإنتاج الفكري لاسترجاع

المعلومات. وقد شملت التطبيقات سلسلة من الوثائق المصممة على أسس النصوص الفائقة حول موضوع النصوص الفائقة بعنوان (Fox, Rous, *Hypertext on Hypertext* and Marchionini, 1991) ومشروع برسيوس Perseus Project ، وهو عبارة عن رحلة تعتمد على النصوص الفائقة في بلاد الإغريق (Bannon, 1991, Myloans, 1992) ، وباثماك PathMAC وهو أداة تستثمر الوسائط الفائقة لمساعدة الطلبة في دراسة المدخل إلى علم الأمراض (Diaz, 1991) .

وعلى الرغم من تعدد مزاياها ، فإن لهذه النظم عيوبها (Berk and Devlin, 1991) فنظراً لطبيعتها غير الخطية ، فإنه من السهل للمستخدم أن يضل سبيله في غابة الروابط. ولعلاج هذا القصور يمد بعض المصممين المستخدم بصور توضيحية للمسار الذي سلكه، يمكن أن تستخدم للاسترشاد بها. هذا بالإضافة إلى أنه نظراً لأن الروابط الفعلية عادة ما تنشأ في أثناء تنفيذ النظام، فإن الارتباطات المنطقية بين النقاط الارتكازية والتي لم يتنبه إليها المصممون من البداية ، لا تتاح للمستخدم لكي ينفذ منها فيما بعد. ومن الممكن حل هذه المشكلة إما بإتاحة الفرصة أمام المستخدم لكي ينشئ روابط أثناء التعامل مع النظام، أو بإضفاء مزيد من الذكاء على النظام بحيث يقوم ببرمجة روابطه بنفسه.

وتنطوي المرحلة الأخيرة في التحول من الأشكال المطبوعة إلى الأشكال الإلكترونية على التخلص كلية من الوثائق الورقية. وأوضح مثال على ذلك أنه قد حدث بعد ظهور الفهارس الإلكترونية أن تخلت كثير من المكتبات تماماً عن فهارسها البطاقية . وإذا ما سائرناه حتى نهايته المنطقية فإن هذا التحول يعني عالماً بلا ورق على الإطلاق ، حيث يتعامل المستخدمون مع نظم المعلومات الإلكترونية دون سواها ، بالنسبة للوثائق الأولية والوثائق الثانوية على السواء. وعلى الرغم من أنه من الممكن لهذا الموقف أن يحدث يوماً ما، فإن هناك كثيراً من الحواجز التي تحول دون حدوثه فعلاً. وهذه الحواجز ليست تقنية وإنما فكرية وسياسية.

بيئة المكتبات ومراكز المعلومات :

لكي نحدد لأي صوب يمكن للمكتبات ومراكز المعلومات أن تكون متجهة الآن ،

فإنه يتعين التعرف على الموقف الراهن . فالتغير يحدث تدريجياً وبشكل مستمر ، ولكن بمعدلات متفاوتة في مختلف المؤسسات ، ويتوقف ذلك على عدد كبير من العوامل . ونسجل فيما يلي عدداً من التعميمات التي تحاول تصوير موقع مراكز المعلومات على عدة مسارات . ويصور الشكل رقم (٨٢) هذه المسارات التي تقابل الفئات الخاصة بالاختزان والتجهيز والنقل.

| الاختزان | التجهيز | النقل | الورقي | المختلط | الإلكتروني |
|--|--|--|---|---|---|
| مقتنيات ورقية من المصادر الأولية والثانوية | البحث في المصادر الثانوية المطبوعة ، والاسترجاع من المقتنيات الأولية الورقية | الإرسال البريدي العادي للوثائق الورقية | المصادر الأولية ورقية، والمصادر الثانوية ورقية أو إلكترونية | البحث على الخط المباشر في المصادر الثانوية، والاسترجاع من المقتنيات الأولية الورقية | إلكترونية تماماً |
| | | | | النقل الرقمي للوثائق الورقية الأصلية | النقل الرقمي عبر شبكات الحاسبات الموزعة |

الشكل رقم (٨٢) : الاختزان والنقل والتجهيز في المكتبات ومراكز المعلومات .

وقد سادت الطرق المعتمدة على الورق دون سواه قبل دخول تقنيات الحاسبات الآلية، حيث كانت مراكز المعلومات تبني وتصون مجموعات ضخمة من الوثائق الأولية التي كانت تختزن أساساً في شكل مطبوعات ورقية، وإن كانت هناك بعض الوثائق، وخاصة تلك المرتبطة بالاهتمامات الأرشفية، التي تختزن في مصغرات فيلمية . وكانت عمليات البحث عن الوثائق الأولية في المجموعات تتم في الفهارس الورقية لمراكز المعلومات، أو في العديد من مصادر التكشيف والاستخلاص التي تشتري من الناشرين التجاريين. وكانت عمليات البحث في الفهارس لا تسترجع سوى تلك الوثائق التي

يقتنيها مركز المعلومات . أما عمليات البحث في أي من الأدوات الأخرى فغالباً ما كانت تكشف عن وثائق غير مقتناة محلياً ، وكان من المتعين طلبها من بعض المصادر الأخرى البعيدة، وعادة ما تكون أحد مراكز المعلومات الأخرى . ونظراً لانكماش ميزانيات مراكز المعلومات، ونمو أعداد المصادر الأولية والثانوية، وتزايد تكلفتها، لم يعد بإمكان أي مركز للمعلومات سوى اقتناء نسبة ضئيلة من الإنتاج الفكري المنشور ، مما يضاعف من أهمية الإمداد بالوثائق عن طريق الإعارة المتبادلة بين المكتبات. إلا أن التأخر لفترات طويلة في الإمداد بالوثائق ، في هذا النظام المعتمد على المطبوعات الورقية دون سواها ، كان أمراً مألوفاً ، لأن الأمر كان يتطلب تصوير كثير من الوثائق أو استنساخها على مصغرات فيلمية ، ثم إرسالها بالبريد .

أما في حالة استخدام مركز المعلومات لكل من المصادر الورقية المطبوعة والمصادر الإلكترونية (الأسلوب المختلط في الشكل رقم (٨٢)) فإنه عادة ما يحتفظ بمقتنياته من المصادر الأولية في شكلها الورقي ، بينما يتم التعامل مع المصادر الثانوية في شكل إلكتروني. ويحدث في كثير من الأحيان أن يكون من الممكن التعامل مع المصادر نفسها في الشكل الإلكتروني أو في الشكل المطبوع في الوقت نفسه. ومن ثم فإنه يمكن لكثير من مراكز المعلومات أن يتوافر بها الفهارس في شكلها الورقي وعلى الخط المباشر في الوقت نفسه، وكذلك الكشافات المطبوعة والتعامل مع الأدوات هذه نفسها في شكلها الإلكتروني عن طريق متعهدي خدمات الاسترجاع مثل دIALOG ، التي يتم التعامل معها من جانب اختصاصيي المعلومات المتفرسين ، وعن طريق الأسطوانات الضوئية المكتنزة المقتناة محلياً ، والتي عادة ما يتعامل معها المستفيدون النهائيون. إلا أن الوصول إلى الوثائق الفعلية التي تعرف بها هذه المصادر ما زال يتم حتماً عن طريق المقتنيات الورقية . ويمكن للإمداد بالوثائق من مراكز المعلومات البعيدة أن يتم بواسطة البريد، أو في حالات متزايدة ، عن طريق التصوير الرقمي عن بعد (الفاكس FAX). (Brown, 1989).

ولم يعمل الأسلوب الثالث، وهو الأسلوب الإلكتروني ، بكامل طاقته بعد . فهو ينطوي على التعامل إلكترونياً مع الوثائق الأولية والمصادر الثانوية ، حيث يمكن إنتاج الوثائق واختزانها إلكترونياً ، كما يمكن الوصول إليها ونقلها عبر مسافات شاسعة عن طريق شبكات الاتصالات بعيدة المدى عالية السرعة . والصورة كما يقدمها الإنتاج الفكري صورة « المكتبة التصويرية » (Mitchell and Saunders, 1991) أو إذا فضلنا

مصطلحاً لا يعبر عن تحيز مؤسسي فهي صورة « بيئة محطة العمل الكونية Universal workstation environment (Dougherty and Hughes, 1991) » وقلماً يكون من الواضح عند مناقشة هذه المفاهيم كم الإنتاج الفكري الأولي الفعلي الذي يمكن أن يكون متاحاً على الخط المباشر ، إلا أن هناك اتفاقاً على أن يكون الهدف هو توفير « ... تسهيلات تستثمر التكامل بين المكتبة والحاسب الآلي ، شفافة واضحة للمستفيد ، فضلاً عن التعامل بشكل شامل مع مرصد البيانات الوطنية على اختلاف أشكالها ، على أن يكون من الممكن الوصول « تصورياً » إلى المعلومات في مكان واحد ، وبحيث تكون محطات العمل في متناول الجميع وبالنسبة لكل المعلومات التي تدعو الحاجة إليها ، هذا بالإضافة إلى منفذ جامع يمكن أن يتعامل مع الوسائط المتعددة على اختلاف أشكالها ، وأن تكون محطة العمل قادرة على الوصول إلى جميع الوسائط ، في أماكن متعددة ، يضاف إلى ذلك إمكان الوصول وبشكل شامل إلى جميع مرصد البيانات بصرف النظر عن موقع المستفيد والمصدر . (Dougherty and Hughes, 1991, P. 11) .

ولهذا السيناريو أو تتابع الأحداث دلالات عميقة متعددة ؛ فمن الممكن لبيئة للمعلومات كتلك البيئة التي عرضنا لها في الاقتباس السابق ، أن تغير الأدوار التي يضطلع بها كل من اختصاصيي المعلومات ومراكز المعلومات كمؤسسات :

١ - لن تظل المكتبات كمبانٍ تضم مجموعات من الوثائق الورقية في الأساس ، محتفظة بأهميتها البالغة ، بل إنها يمكن في الواقع أن تختفي تماماً .

٢ - يمكن للمجموعات أن تصبح فعلاً جميع المصادر المتوافرة في مختلف المواقع التي يمكن التعامل معها محلياً وعن بعد .

٣ - لن يظل اختصاصي المعلومات يقوم في الأساس بدور الوسيط أو همزة الوصل بين المستفيد ومجموعات الوثائق الورقية ، وإنما يمكن أن يساعد في تصميم نظم البحث والاسترجاع القوية ، وفي استخدام مثل هذه النظم لضمان المزيد من الاستقلال للمستفيد النهائي .

ولدينا فعلاً الدليل المؤيد لكل النقاط الواردة في السيناريو الذي عرضنا له توأ .

١ - هناك بعض المكتبات التي تتيح على الخط المباشر النصوص الكاملة لمقالات الدوريات ، لأغراض البحث والاسترجاع والتسجيل من جانب المستفيدين .

٢ - أصبح من الممكن الآن، وبشكل متزايد، للمستفيدين واختصاصيي المعلومات التعامل مع الفهارس ومراصد البيانات الموجودة في مواقع نائية، عن طريق الشبكات، كالإنترنت مثلاً.

٣ - تتزايد أهمية دور المستفيد النهائي، وخاصة في الإفادة من الفهارس المتاحة على الخط المباشر، وغيرها من مراصد البيانات المسجلة على أسطوانات ضوئية مكنزة.

وقد أصبح من الممكن الآن اختزان كميات هائلة من المعلومات بشكل اقتصادي (فالأسطوانة الضوئية المكنزة الصغيرة الواحدة تتسع الآن لاختزان حوالي ٥٠٠ مليون تمثيلة أو حرف) كما أصبح من الممكن أيضاً نقل كميات هائلة من المعلومات عبر الشبكات عالية السرعة في كسور الثانية. ومن ثم فقد توافر العتاد اللازم لاختزان كميات هائلة من الإنتاج الفكري العالمي، وذلك بنصوصه الكاملة مع ما يصاحبها من رسومات بيانية وخرائط وصور. هذا بالإضافة إلى أنه من المتوقع أن تصبح الحاسبات الآلية في منتصف العقد الأخير من القرن العشرين أسرع من تلك التي توافرت في نهاية العقد التاسع من القرن نفسه، بمقدار ألف مرة، وأن يكون بمقدور الشبكة القومية للبحوث والتعليم (NREN) نقل بلايين التمثيلات في الثانية، وذلك في نطاق بيئة موزعة متزايدة الترابط لاستخدام الحاسبات الآلية.

إلا أنه على الرغم من قدرة عتاد الحاسبات الآلية على مساندة اختزان ونقل كميات هائلة من المعلومات، هناك عدد من الحواجز التي لا يستهان بها، تقف حجر عثرة في سبيل توفير كل هذه المعلومات والإفادة منها بشكل فعال. وتنطوي إحدى هذه العقبات على الالتزام بقانون حقوق التأليف والنشر. فكما يرى فيبر (1990) Weber فإنه :

« على الرغم من أن قانون حقوق التأليف والنشر يعتبر حماية مهمة لكل من المؤلفين والناشرين، فإنه بنهاية هذا القرن سوف يصبح عاطلاً إلى حد بعيد، في الحالات التي يتم فيها إنتاج المعلومات وتوزيعها وتبادلها إلكترونياً. وسوف يؤدي هذا الموقف إلى عزوف المؤلفين والناشرين عن التخلي عن النشر التقليدي على الورق ». (P. 80)

وعلى الرغم من أن الناشرين سوف يلجأون حتماً لاتخاذ التدابير الفنية لمراقبة التعامل مع الوثائق وتحصيل حقوق المؤلفين، فإن هذه المهمة تبدو معقدة، وخاصة إذا أخذنا في الحسبان صعوبة تتبع مسار الوثائق في انتقالها من شخص إلى آخر. ومن ثم،

فإنه على الرغم من أن النشر الإلكتروني يعد إحدى الصناعات النامية ، فقد أدت التزامات قانون حقوق التأليف والنشر إلى الحد من سرعة تبنيه .

وتتصل إحدى العقبات الأخرى التي تعترض سبيل نظام المعلومات الإلكتروني في سدهاء ولحمته ، تتصل مباشرة بكثير من المواد الواردة في هذا الكتاب ، وتنطوي على قضايا التعامل مع المعلومات ومصادر المعلومات . ومن بين مشكلات الوصول إلى المعلومات تلك المشكلة الناتجة عن الارتباط المتزايد عبر الشبكات ، والذي سرعان ما يؤدي إلى تزايد عبء المعلومات على نطاق يتجاوز الحدود بشكل غير عادي. فقد أصبح اختيار المعلومات المناسبة والمفيدة ، وهو المشكلة المحورية لاسترجاع المعلومات ، أكثر صعوبة في البيئة المتشابكة التي تتكون من كميات متزايدة من الاتصالات الرسمية وغير الرسمية ، والتي لم يخضع كثير منها بعد للتنظيم وفقاً لأية طريقة نسقية^(١).

وتنطبق الملاحظات نفسها المتعلقة بالافتقار إلى التنظيم الفعال على التعامل مع النصوص الكاملة للوثائق . فالبحث بالكلمات المفتاحية للأعداد المتزايدة من هذه الوثائق عادة ما يؤدي إلى استرجاع مجموعات من المعلومات غير المناسبة في معظم الأحيان ، والتي تحول ضخامة حجمها دون تصفحها بشكل فعال. وقد نجحت بعض النظم في تحسين هذا الموقف بتطوير برمجيات لترتيب الوثائق طبقاً وفقاً لاحتمال صلاحيتها ، إلا أن المشكلة الرئيسة هي مشكلة التحليل الموضوعي والتعبير عن ناتج هذا التحليل ؛ فالنصوص الكاملة للوثائق ليست مكثفة أو مركزة في المحتوى كغيرها من أشكال التعبير كالمستخلصات والمصطلحات الكشفية. هذا بالإضافة إلى أن أساليب الخط المباشر ، والتي لا غنى عنها لوصف وتنظيم واسترجاع الصور والأصوات وغيرها من المعلومات غير الهجائية الرقمية ، ما زالت في مراحل الاستكشاف المبكرة.

(١) هناك العديد من المحاولات الجارية لتنظيم المصادر المتاحة عن طريق الإنترنت، ومن بين هذه المحاولات بحث يقوم به مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC لوضع نظام لفهرسة هذه المصادر . وينطوي بحث هذا المركز على دراسة لدى إمكان إجراء الفهرسة آلياً (1991م ، OCLC)

وهكذا ، فإن التحديات الجوهرية لاسترجاع المعلومات تشمل ما يلي :

١ - إتاحة فرصة التعامل : فبإمكان المستفيدين الآن التعامل مع أعداد من الملفات المتاحة على الخط المباشر ، أكثر مما توافر لهم من قبل ، وتشتمل هذه الملفات على أشكال مختلفة من طرق التعبير كالنصوص والصور والأصوات . فكيف يمكن توفير الأدوات الفعالة لمراقبة الشبكة بشكل مستمر ، بحثاً عن المعلومات المناسبة بناء على سمات اهتمامات المستفيدين ؟ أضف إلى ذلك ، كيف يمكن تصميم واجهات التعامل الذكية القوية ، والتي يمكن أن تساعد المستفيدين في تتبع المعلومات المتاحة في أماكن متعددة ، والمقدمة في شكل نصوص وأرقام وصور وأصوات ، والبحث عن هذه المعلومات واستعراضها أو تصفحها ؟

٢ - الإمداد بالوثائق : بإمكان المستفيدين الآن الحصول على أعداد كبيرة من الإشارات الوراقية على الخط المباشر ، إلا أن الحصول على الوثائق الأولية ما زال معتمداً على الورق في معظم الأحيان . كيف يمكن التغلب على قضايا حقوق التأليف والنشر ، بحيث يصبح من الممكن الحصول على جميع المعلومات إلكترونياً ؟

آراء أخرى حول المستقبل :

بينما كان هذا الكتاب في مرحلة التأليف ، كان أحد مؤلفيه يقوم أيضاً بتحرير كتاب يضم عدداً من المقالات حول مستقبل المكتبات (Lancaster, 1993 b) . ويتكون هذا الكتاب من إسهامات عدد من المبرزين من المكتبيين وعلماء المعلومات وغيرهم من العلماء الباحثين من الولايات المتحدة الأمريكية وغيرها (*) . ويرى المؤلفون وبالإجماع تقريباً أنه من الممكن النظر إلى التطورات التقنية وغيرها من التغيرات التي تحدث في العالم من حولنا ، إما باعتبارها تحدياً للمكتبات وإما باعتبارها فرصة نادرة لمهنة المكتبات لكي تدعم قيمتها وتؤكد أهميتها للمجتمع أكثر مما كانت عليه من قبل . ويمثل المشاركون في ذلك الكتاب أنواعاً كثيرة من المكتبات وما يتصل بها من المؤسسات في أربع قارات . وفي مثل هذا التنوع في الخلفيات والمنطلقات يمكن للمرء أن يتوقع تفاوتاً كبيراً في الرؤى والتصورات ، إلا أنه قد تبين في الواقع وجود اتفاق ملحوظ في غالب الأحيان .

(*) عرضنا هذا الكتاب في : دراسات عربية في المكتبات وعلم المعلومات (كتاب دوري محكم) ع ١ ، ص ١ : يناير

١٩٩٦ . ص من : ٢٠٥ - ٢١٦ . (المترجم) .

فها هو ذا بنيمان (Penniman (1993 يؤكد على ضرورة أن تكون المكتبات نشطة فعالة لا سلبية خاملة ، وأن تحرص على إيصال المعلومات لا على اختزانها . كما يؤكد أيضاً على ضرورة تقييم المكتبات على أساس ما تقدمه من خدمات لا بناء على ما تقتنيه من أرصدة . ومن موقعه كرئيس لمجلس الموارد المكتبية Council on Library Resources يرى بنيمان أنه يتعين على اختصاصيي استرجاع المعلومات تركيز طاقاتهم على المشاركة في البحث ، تلك المشاركة التي تضمن للمكتبات أن تصبح هي نظم الإمداد بالمعلومات ، تلك النظم التي يحتاجها المستقبل ، وأن تتوافر للمهنة المهارات القيادية اللازمة لجعل المكتبات نظاماً للإمداد لا غنى عنها . ويرى بنيمان أن المكتبات الآن في خطر نظراً لافتقادها القيادة الحقيقية.

وفي حديثها عن بيئة المكتبات الأكاديمية تركز مولهولت (Molholt (1993 على التطورات التكنولوجية ، وهي ترى أن تزايد قدرات تقنيات المعلومات وما تحرزه هذه التقنيات من تقدم ، ذلك التقدم الذي يشكل تهديداً لوجود المكتبات « التقليدية » يمكن أن يؤدي في المستقبل إلى تعزيز مكانة المكتبات كمؤسسات والمكتبات كمهنة ، بشكل لم يتحقق من قبل . وتتفق مولهولت مع بنيمان على الحاجة إلى حدوث تغيرات في المهنة :

« سوف تؤدي تقنيات الشبكة وموارد المعلومات الموزعة ، وتقنيات البرامج غير الخطية التي تكفل الترابط الوثيق بين موارد المعلومات ، وكذلك الاتجاه نحو النظم التفاعلية التي تقدم المعلومات بالصوت والصورة ، إلى تغير صورة المكتبات وما يضطلع به المكتبيون . فنحن بحاجة إلى التخطيط لإحداث تغيير في التركيز ، بحيث يتحول المكتبيون والمكتبات من سدنة للكتب إلى أدلة مرشدة في عالم المعرفة ... وسوف تكون قدرة المكتبيين على التخيل وعلى الخروج عن نطاق المكتبة إلى آفاق المعلومات الرحبة ، هي معايير الحكم على مكتبيي المستقبل . »

أما داوولن (Dowlin (1993 الذي يمثل المكتبات العامة ، ويانج (Young (1993 الذي يمثل المكتبات الأكاديمية فيتفقان مع بنيمان ومولهولت ، حيث يؤكدان أيضاً الحاجة إلى إحداث تغيير في المهنة . وربما اختلفت أساليب التعبير التي استعملها هؤلاء المؤلفون ، إلا أن الرسالة واحدة في الأساس . فداوولن ينظر إلى المكتبة في صورتها الراهنة كقلعة ، بينما ينبغي أن تكون في الواقع خطأً للأنايب . وكما هو الحال بالنسبة لبنيمان ، فإنه يرى الخطر كامناً في عجز المهنة عن تقديم القادة الذين تحتاج إليهم لتحويل المكتبة

من قلعة إلى خط للأنايب ؛ فمعاهد المكتبات تؤهل خريجين في مستوى الحرفيين المتمرسين أو مقاولي التنفيذ ، ولكنها لا تؤهل المهندسين المعماريين ذوي الرأي والرؤية. هذا ، ويبني داوولن رؤيته لمستقبل المكتبة العامة اعتماداً على الخطط الخاصة بتطوير مكتبته ، وهي مكتبة سان فرانسيسكو العامة . فهذه المكتبة لن تكون بأي حال تجسيدا لفكرة « المكتبة بلا جدران » التي تنبأ بها بعض الكتاب (Lancaster, 1982) ، ولكنها ستكون مبنى « ذكياً » فعلاً ، يشتمل على مقومات (ستوديوهات) البث السمعي والبصري ، القادرة على إيصال الخدمات المكتبية إلى المنازل . ويتوقع داوولن أنه بحلول عام ٢٠٠٠ سوف تكون جميع منازل سان فرانسيسكو مرتبطة بالمكتبة بنظام التلفزة السلكية أو بخطوط الهاتف ، وأن مبنى المكتبة سيكون هو « النقطة الارتكازية للشبكة المجتمعية » . كما يعتقد أيضاً أنه بإمكان التقنيات الإلكترونية أن توفر للمستفيدين المكتبة التي تكتسب « طابع وروح مجتمع المدينة الصغيرة » والتي تكفل في الوقت نفسه مقومات الاتصالات الكونية الفورية.

ويؤكد يانج (Young (1993 أنه سوف يتعين على المكتبات توفير المعلومات بأكثر الأشكال ملائمة للمستفيد ، سواء كان الشكل المطبوع أو الإلكتروني أو الضوئي ، أو أي شكل آخر يمكن أن يظهر في المستقبل . (ومن الجدير بالملاحظة أن معظم من شاركوا في كتابة فصول ذلك الكتاب لا يرون في أشكال الأوعية الجديدة بديلاً عن القديمة ، وإنما مصاحبة لها) . ويعتقد يانج أن الهدف النهائي للمهنة هو إنشاء « المكتبة التخيلية أو التصورية » التي تتكون من « إجمالي مجموع المعلومات التي يمكن الوصول إليها والمتاحة في أي مكان » . ووفقاً لهذا التصور يصبح مقر المكتبة مجرد « نقطة ارتكازية للاسترجاع » .

ويستعرض رائت (Raitt (1993 تقنيات المعلومات المتاحة في الوقت الراهن وتلك التي يمكن أن تتاح في المستقبل القريب ، ويرى أنه ينبغي أن يكون من بين الأدوار المهمة التي ينبغي أن يضطلع بها المكتبي تقييم التقنيات الملائمة وتزويد الآخرين بالمعلومات حول قدرات هذه التقنيات . فالتقنيات في الوقت الراهن لا يطورها المسئولون عن بث المعلومات ، وإنما تطورها « صناعة معلومات موازية » (وهي صناعات الحاسبات

ووسائل الترويج في الأساس) . ويقصد رائت القول بأن المكتبيين بحاجة لأن يكونوا أكثر إيجابية في تحديد كيفية الاستفادة من التقنيات.

ويتفق لايين (1993) Line الذي يمثل النظرة البريطانية مع بنيمان على أنه ينبغي أن يتم تقييم المكتبات على أساس ما تقدمه من خدمات لا بناءً على ما تملكه من مقتنيات. ويرى لايين أن التقنيات سوف تحد من كون المكتبة « مكاناً يؤمه المستفيدون وتدعم من مقوماتها كمورد يمكن الاستفادة منه عن بعد » . ويتفق معظم المؤلفين فعلاً على أن المكتبات كمؤسسات وكذلك مقار المكتبات ومبانيها سوف تظل قائمة فعلاً ، إلا أن الدور الذي تضطلع به سوف يتغير.

ويمثل الدول النامية كل من نيلاميفان (1993) Neelameghan وكريمر (1993) Kremer، ويعتقد الأول أنه سيكون للتقنيات آثارها الإيجابية على دول العالم الثالث ، حيث سيكون بمقدور المكتبات في هذه الدول « توفير فرص الوصول إلى موارد للمعلومات أكثر اتساعاً ، عن طريق الشبكات على المستويات القومية والإقليمية والدولية ، وكذلك الاستفادة من مراصد البيانات على الوسائط الضوئية الإلكترونية ، كالاسطوانات الضوئية المكتنزة CD-ROM » . أما كريمر التي تركز على الموقف في البرازيل ، فإنها على الرغم من عدم التهوين من قدر الدور المحتمل للتقنيات ، تؤكد على أن المكتبات ومرافق المعلومات ينبغي أن تتطور في تلك الدولة إلى الحد الذي تستطيع به تأكيد أهميتها في تلبية الاحتياجات الاجتماعية والتقنية والثقافية للأمة . وبعبارة مختلفة إلى حد ما، فإنه يمكن للتقنيات مساعدة مكتبات العالم الثالث في تحقيق أهدافها، إلا أن استخدام التقنيات الحديثة لا ينبغي أن يكون هدفاً في حد ذاته .

أما نظرة كلجور (1993) Kilgour إلى مكتبة المستقبل فهي أقل تقليدية من نظرة كثير من المشاركين الآخرين ، حيث يضع تصوراً لنظام مكتبي إلكتروني يتكون من مرصد بيانات مركزي يضم النصوص الكاملة للكتب والمقالات وغيرها من أوعية المعلومات، بالإضافة إلى مراصد بيانات تكفل للمكتبات التعامل معها عن بعد ، تضم مختلف كشافات هذه النصوص الكاملة .

وربما كان من المفارقات الغربية أن يأتي أكثر التصورات مبالغة في الابتعاد عن الطابع الحالي للمكتبات ، من جانب أحد الباحثين المستفيدين من المكتبات (لورين سايلر Lauren Seiler) لا من جانب مكتبي ، على الرغم من أن سايلر يساعده ويستحثه أحد أساتذة المكتبات وعلم المعلومات (توم سير برينانت Tom Surprenant) . ويرى كل من سايلر وسير برينانت ، عن ثقة أننا نكاد نرى نهاية مكتبة المطبوعات ، وأن الطباعة على الورق في سبيلها للانقراض ، ويقدمان وصفاً لعالم من المكتبات جميع مصادر المعلومات والإبداع والترويج فيه إلكترونية ، كما يريان أن « مركز المعلومات التصوري » الذي يقصده يانج (1993) Young ربما يكون هو الهدف النهائي لمهنة المكتبات . وفي مركز المعلومات التصوري يحل محل أرفف الكتب صور هذه الأرفف ، وبإمكان كل مستفيد أن يتجول إلكترونياً في جميع أرجاء مستودع العالم من المعرفة المسجلة دون مغادرة منزله .

دعنا ننقل الآن إلى مجموعة أخرى من التنبؤات الواردة في كتاب *Information UK 2000* الذي تولى تحريره كل من مارتن Martyn وفيكركز Vickers وفيني Feeney (1990) . ويقدم هذا الكتاب نتائج دراسة ، أجريت بمبادرة من المكتبة البريطانية ، للتنبؤ بالطرق المحتملة لإنتاج المعلومات واختزانها وتداولها خلال العقد القادم أو في هذه الحدود . وقد تعرضت الدراسة لما هو أكثر من المنظور التقني ، حيث تناولت القضايا الاجتماعية والاقتصادية وما يتصل بها . وقد بدأ التخطيط للمشروع في مطلع عام ١٩٨٩ م ، حيث تجمعت معظم البيانات في نهاية العام نفسه وبداية عام ١٩٩٠ م . وتستند الدراسة إلى مدخلات أمكن الحصول عليها من حوالي ستين « خبيراً موضوعياً » تنتظمهم سلسلة من إحدى عشرة مجموعة عمل . وقد قام كل عضو من أعضاء فرق العمل هذه بإعداد تصوره المستقبلي مستقلاً عن الآخرين ، حيث تم بعد ذلك دمج هذه التصورات معاً لتشكيل تصوراً موحداً لكل مجال على حدة . كما تم أيضاً الحصول على مساهمات من أفراد لم يكونوا ضمن فرق العمل هذه .

ويتكون أحد عشر فصلاً من فصول الكتاب الخمسة عشر من تقارير مجموعات العمل التي تغطي المجالات التالية : الاتجاهات الاجتماعية ، والتقنيات ، ودور المحفوظات والمكتبات ومرافق المعلومات (أقل تقارير لجان العمل حجماً ، مما يمكن أن يدل على أن المكتبات ومرافق المعلومات التقليدية لم تكن تعتبر على قدر كبير من الأهمية

في الصورة العامة) ، والتسجيل والاستنساخ ، والبنية الأساسية للاتصالات ، والنشر وأشكال الأوعية الجديدة والتوزيع والتسويق ، وأوجه الاستفادة الفردية والمحلية من المعلومات ، والمنظمات واستثمارها للمعلومات ، والقوى البشرية والتأهيل والتدريب ، وقضايا المستفيدين من المعلومات ، وقضايا السياسة من وجهة نظر واضعي السياسات.

ويشتمل كتاب *Information UK 2000* على عدد كبير جداً من التوقعات المستقبلية وخاصة في مجالات التقنيات ، إلى الحد الذي يحول دون تقديم تلخيص فعال لهذه التصورات . ونظراً لأن موضوع كتابنا هذا « أساسيات استرجاع المعلومات » هو نظم استرجاع المعلومات ، فإن هذه المناقشة تركز على التوقعات المستقبلية التي تتصل بالمكتبات ومستقبل مجال المكتبات وعلم المعلومات . ومن بين التنبؤات أن تؤدي ضغوط الميزانيات إلى التحول من شراء المواد إلى شراء حق الوصول إلى المعلومات ، وأن ينشأ « تجمع ضخمة » يضم المكتبات الكبرى ليكون أساساً للمقتنيات التي يمكن أن تتاح للمكتبات الأخرى إمكانية الوصول إليها ، وأن تتحول المكتبات المحلية إلى مراكز للتوجيه في المقام الأول . وليس هناك في هذه التنبؤات ما هو مروع إلى حد بعيد ، وإنما يمكن في الواقع رد هذا النوع من التنبؤات إلى حوالي عشرين عاماً مضت . أما التنبؤ المثير للجدل أكثر من غيره فهو الخاص بأقول نجم المكتبات الأكاديمية بوجه عام ؛ فوفقاً لما ذهب إليه المحررون ، سوف تعمل الأقسام الأكاديمية على تدبير مقومات تعاملها مع المصادر الإلكترونية ، أما المكتبة الأكاديمية فسوف تصبح مجرد قاعة للدرس لا أكثر.

أما فيما يتعلق بالتأهيل والتدريب فإن التنبؤات تشير إلى تضاؤل فرص التوظيف في المؤسسات « التقليدية » (أي المكتبات) إلا أن أنواعاً جديدة من الوظائف سوف تنشأ ، وهذه تشمل :

« البحث وتحليل المعلومات لدعم اتخاذ القرارات ووضع السياسات ، ومعالجة المعلومات ، بما يتراوح بين صياغة التقارير الفنية ، والاضطلاع بمسئولية إدارة ومعالجة أنشطة إنتاج المعلومات برمتها في إحدى المؤسسات ... وسوف تكون هناك زيادة في أعداد المستشارين والوسطاء » . (P. 29) . وتشير التنبؤات أيضاً إلى حدوث تحول عن التدريب الرسمي أو المؤسسي في المكتبات وعلم المعلومات ، وأن تهتم البرامج

الدراسية في المستقبل بوجه خاص بتقنيات المعلومات والإدارة ، ثم « بالجوانب الاجتماعية » ولكن ليس بالقدر نفسه .

ويرى المحررون أن مهنة المكتبات والمعلومات ، كما هي الآن ، عرضة للتهديد :

« فالتوسع في استخدام التقنيات ، وزيادة الاعتماد على الأساليب الآلية ، وتزايد تآلف المجتمع بأسره مع الحاسبات الآلية ، وتزايد حدة الضغوط التجارية من أجل بيع مختلف أنواع منتجات المعلومات للجمهور ، كل هذه العوامل سوف تؤدي إلى إضعاف موقف المكتبيين التقليديين أو علماء المعلومات » (P. 262) .

ولا أحد يعرف بالطبع أي التصورات التي طرحت حتى الآن من جانب المكتبيين وغيرهم يمكن أن تكون أقرب إلى الصحة . وهذا أمر في حد ذاته لا يهم ، أما المهم فعلاً فهو حدوث تغيرات جوهرية في المستقبل ، في كيفية إتاحة مصادر المعلومات والإبداع والترويج ، كما أن هذه التغيرات سوف يكون لها أثرها العميق في مهنة المكتبات والمعلومات .

وهناك ولا شك كثير من التغيرات التي تحدث الآن في البيئة التي تعمل بها المكتبات ، كما أن مستقبل المكتبات سوف تحكمه وإلى حد بعيد التطورات التي تحدث خارج نطاق السيطرة المباشرة لهذه المؤسسات . وأبرز العوامل المؤثرة في مستقبل المكتبات هو ما سيحدث في صناعة النشر وفي نظام الاتصال العلمي برمته . فمن الواضح أننا سوف نشهد نشر المزيد والمزيد من مصادر المعلومات في شكل إلكتروني ، حيث تتضاءل أهمية الطباعة التقليدية على الورق . أما ما لا نراه بالقدر نفسه من الوضوح فهو كيفية إتاحة مصادر المعلومات الإلكترونية هذه في المستقبل ؛ فهل سيتم الوصول إليها عبر الشبكات الإلكترونية ، أم أنها ستوزع كمنتجات مادية كالأسطوانات الضوئية المكتنزة أو غيرها من الأشكال التي يمكن أن تظهر في المستقبل . ومن الممكن للأشكال الإلكترونية التي يتم توزيعها أن تقتنى من جانب المكتبات بالطريقة نفسها التي يتم بها اقتناء المصادر المطبوعة ، أما المصادر التي لا يمكن الوصول إليها إلا عبر الشبكات الإلكترونية فتمثل للمكتبات مجموعة مختلفة من المشكلات ، وتفرض عليها دوراً مختلفاً إلى حد ما باعتبارها مرافق للإمداد بالمعلومات .

كذلك سوف يتحدد شكل مكتبة المستقبل تبعاً لشكل المؤسسة التي تتبعها : فالولايات المتحدة تتجه بسرعة نحو « الجامعة الإلكترونية » ، التي يتم فيها الجانب الأكبر من التواصل للأغراض التعليمية والعلمية والبحثية عن طريق الوسائط الإلكترونية . وقد ناقش دافيد لويس David Lewis ، أحد المكتبيين بجامعة كولومبيا الانعكاسات المحتملة لهذا الاتجاه على المكتبة الأكاديمية في المستقبل (Lewis, 1988) . وهو يرى أن أتمتة النظم القديمة لا تكفي ، وإنما سوف يتطلب الأمر إعادة تنظيم شاملة للمؤسسات . هذا بالإضافة إلى أن المستفيدين سوف يعتمدون على المكتبات أكثر مما كانوا عليه من قبل :

« فمن الممكن للطلاب أن يتوقعوا من المكتبة أن تكون قوية ، يسهل التعامل معها كما هو الحال بالنسبة للوسائط التعليمية الإلكترونية . ونادراً ما يكون من السهل . للأسف ، التعامل مع المكتبات . وإذا ما أصبح التحليل بواسطة الأدوات الإلكترونية الحديثة أيسر وأكثر فعالية من البحث في المكتبات ، فإنه من المتوقع أن يستعين الطلاب بهذه الأدوات الحديثة لا بالمكتبات . وما لم ترتفع المكتبات بمستوى خدماتها بحيث تظل مرافق تعليمية أساسية ، فإنها عرضة لأن تصبح غير صالحة للعملية التعليمية . وإذا ما أدت الظروف إلى حدوث ذلك فعلاً ، فإنه من السهل التنبؤ بحدوث انخفاض في مخصصات تمويل المكتبات » (P. 293) .

الذكاء الاصطناعي والنظم الخبيرة :

من المظاهر التقنية التي ينظر إليها غالباً باعتبارها حلاً لجميع مشكلات المكتبات ومرافق المعلومات ، الذكاء الاصطناعي (A I) أو النظم الخبيرة . وعادة ما تصطدم أية مناقشة لتطبيقات الذكاء الاصطناعي بافتقار المصطلح للتعريف الذي يحظى بإجماع القبول . ولم تبذل عدة كتب مؤلفة حول الذكاء الاصطناعي أية محاولة حقيقية فعلاً لتعريف هذا المصطلح ^(١) . والأدهى من ذلك أن المصطلح يستعمل بلا حرص للدلالة غالباً على عمليات لا دخل فيها للذكاء الآلي (كالاختيار مثلاً من جانب البشر من بين محتويات قائمة اختيار يعرضها الحاسب الآلي) . وإنه لمن العجيب حقاً أن يرى البعض

(١) حرص كورتسفايل Kurzweil على توضيح أهداف الذكاء الاصطناعي . اقتباساً من إيلين ريث Elaine Rich ، وهي « كيف نجعل الحاسبات قادرة على تنفيذ أعمال ينجزها البشر الآن بشكل أفضل » (P. 69) .

النظر إلى جهود التكشيف أو الاستخلاص الآلي ، كتلك التي قام بها الرواد من أمثال لون (1958) Luhn وباكسنديل (1958) Baxendale حيث كان يتم التقاط الكلمات من النص بناء على معايير مدى التردد ، والتي كانت تعد وقتئذ من تطبيقات علم اللغة الحاسبي ، النظر إلى هذه الجهود باعتبارها من تطبيقات الذكاء الآلي !.

وربما يكون فنلي (1992) Fenly قد قدم أوضح وأوجز بيان لماهية الذكاء الاصطناعي أو ما يمكن أن يكون فعلاً :

« لقد أمكن تصميم برامج الحاسبات الآلية القادرة على الاستنتاج كالبشر ، والتي يمكن أن تكون قادرة على التعلم من أخطائها ، والتي تقوم ، بسرعة ومهارة ، بإنجاز مهام ، عادة ما يضطلع بها خبراء من البشر يتسمون بالندرة وارتفاع التكلفة » (P. 52).

وبعبارة أخرى ، فإن الذكاء الاصطناعي يحاول تطوير نظم قادرة على تنفيذ بعض المهام التي يضطلع بها عادة خبراء متخصصون في مجالات معينة ، وربما كان أوضح مثال على ذلك التشخيص الطبي . ولهذا السبب ، فإن مثل هذه النظم غالباً ما تسمى « بالنظم الخبيرة » . ويستعمل الآن كل من المصطلحين « النظم القائمة على المعرفة » و« النظم القائمة على القاعدة » تبادلياً إلى حد ما مع « النظم الخبيرة » . والسبب في ذلك أن هذه النوعية من النظم ينبغي أن تُعطى جانباً من المعرفة (كالأعراض والعلاقات المرتبطة بحالة مرضية معينة مثلاً) لكي تعتمد عليه في عملها ، وتتكون بعض الارصدة المعرفية هذه من القواعد ، كقواعد الفهرسة الوصفية مثلاً .

والواقع أنه نظراً لاستناد الفهرسة الوصفية إلى القواعد ، فإنها تأتي في مقدمة المجالات المرشحة لاستخدام النظم الخبيرة . وقد بذلت بعض الجهود فعلاً في هذا المجال (منها على سبيل المثال (Weibel, 1992; Borko and Ercegovac, 1989; Jeng, 1986; and Schwarz, 1986) إلا أن فنلي (1992) Fenly يرى أن النتائج التي تحققت حتى الآن تبدو غير مقنعة ، ويعتقد أن نظام الفهرسة الذي ينطوي على خبرة حقيقية أصعب في تنفيذه بكثير من النظام الذي يكتفي بمجرد عرض قواعد الفهرسة في شكل آلي . وكما يسجل فنلي أيضاً ، فإن « النظم الخبيرة الحقيقية ، التي تتمتع بالعمق والقوة اللازمين لحل المشكلات الجوهرية العويصة ، يتطلب تطويرها وقتاً طويلاً فضلاً عن ارتفاع التكلفة » (P. 54).

إلا أنه يرى في مقابل ذلك أنه من الممكن أن تكون هناك مشكلات معينة في الفهرسة الوصفية، يتطلب حلها قدراً غير عادي من الجهد الفكري، وأنه يمكن لهذا النوع من المشكلات أن يبرر تكلفة تطوير النظم الخبيرة التي تغطي المجال برمته. ومن بين هذا النوع من التطبيقات فهرسة السلاسل.

هذا، وقد أجرى وايبيل (Weibel (1992 بحثاً في مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC حول جدوى الفهرسة الوصفية الآلية المعتمدة على صور صفحات العنوان، إلا أنه يرى « خيطاً من مجافاة الحقيقة » في كثير من البحوث التي أجريت في هذا المجال. كما يرى أن هناك « عقبات ضخمة في سبيل تطوير نظم الإنتاج » وأنه لا يمكن لأساليب النظم الخبيرة « أن تغير من أساليب التجهيز الفني في المكتبات في السنوات الخمس القادمة » (P. 72). إلا أنه يعتقد، شأنه في ذلك شأن فني، أنه من الممكن لمهام تخصصية معينة في الفهرسة أن تفيد من استخدام النظم الخبيرة، و«تصميم» الطرق الآلية للفهرسة « بذكاء » أهم في نظره من كون هذه الطرق تنقسم بالذكاء.

وتعيين المصطلحات للوثائق للدلالة على الموضوعات التي تتناولها، نشاط آخر يمكن أن يفيد من تطبيقات الذكاء الاصطناعي. وعلى الرغم من أن الكشف الموضوعي لا يمكن أن يكون مستنداً إلى القواعد كالفهرسة الوصفية، فإنه لا مناص من اتباع قواعد معينة. وفي النظم باللغة الضخامة، كتلك التي تديرها المكتبة القومية للطب، عادة ما تكون هذه القواعد كثيفة. فمن الممكن على سبيل المثال لإحدى مجموعات القواعد أن تحدد أي الرعوس الفرعية يمكن أن يستعمل مع أي الفئات من الرعوس الرئيسية.

وفي المكتبة القومية للطب يتم الآن وضع برنامج تفاعلي باسم Med Ind Ex، بناء على أسس الذكاء الاصطناعي، لمساعدة المكشفين في استعمال الـ *Medical Subject Headings* للتعبير عن المحتوى الموضوعي للمقالات البيوطبية (Humphrey, 1992). وبإمكان هذا البرنامج في الأساس إنجاز مهمتين رئيسيتين :

١ - بإمكانه مساعدة المكشف على سرعة تعيين مصطلح معين أو نوعية معينة من المصطلحات.

٢ - بإمكانه تصحيح أداء المكشف عندما يستعمل مصطلحاً غير مناسب.

وتدعي بعض الطرق الأخرى المستخدمة في الكشف الآلي أو الكشف بمساعدة الحاسب الآلي والتي حظيت بالوصف في الإنتاج الفكري ، استخدام أساليب الذكاء الاصطناعي أو أساليب النظم الخبيرة ، إلا أنه من الصعب أن نفهم كيف يمكن للنظم التي تقوم بتعيين المصطلحات للوثائق ، بناء على التشابه بين الكلمات التي ترد في نص الوثيقة (كالعناوين والمستخلصات على سبيل المثال) وفي « سمات » الكلمات المرتبطة بالمصطلحات ، أن ينظر إليها باعتبارها تنطوي على ذكاء اصطناعي . ويمكن من ناحية أخرى للذكاء الاصطناعي أن يكون له دور في هذا المجال إذا أصبح بإمكان نظام الكشف أن يتعلم من أخطائه ، وبذلك يكون قادراً على الارتفاع بمستوى أدائه .

هذا ، وقد بذلت جهود كثيرة في تصميم وتطوير « واجهات التعامل الذكية » التي تساعد في الإفادة من مرصد البيانات عن طريق شبكات الخط المباشر . فقد قام هيو (1987) Hu على سبيل المثال بتقييم إحدى هذه الواجهات التي صممت لمساعدة المستفيد في اختيار مرصد البيانات الذي يمكن أن يكون أكثر ملاءمة من غيره بالنسبة إلى حاجة بعينها إلى المعلومات . وقد تبين من دراسة هيو أن هذه الواجهة على وجه التحديد تعمل معتمدة وبشكل كلي تقريباً على استخدام قوائم الاختيار والتي ينتقي منها المستفيد ، ولا تبدي أي دليل على « ذكاء » حقيقي في اختيار مرصد البيانات .

وقد تم تصميم واجهات تعامل أخرى لمساعدة المستفيد في صياغة استراتيجية البحث التي تعبر عن الحاجة إلى المعلومات بالشكل الملائم . وقد حظي العديد من هذه الواجهات بالدراسة الوصفية التحليلية من جانب كل من فيكري (1992) Vickery وألبريكو وميكو (1990) Alberico and Micco . وبعض هذه الواجهات يعمل بقوائم الاختيار ، وبعضها يحث المستفيد بتوجيه أسئلة مصممة لتحديد مجال عملية البحث بشكل مفيد ، والبعض الآخر يتلقى المدخلات من المستفيد في شكل تعبير سردي عن الحاجة إلى المعلومات . وعلى الرغم من أن كثيراً من هذه الواجهات أدوات واضحة ومفيدة ، فإنه من غير الواضح ما إذا كانت تنطوي فعلاً على استخدام الذكاء الاصطناعي أم لا .

ولم تتحقق بعد الأداة الشاملة للرد على الاستفسارات التي تصورها دانا (1916) Dana ، إلا أنه قد تحقق ولا شك قدر من التقدم نحو تطوير النظم التي يمكن أن ترشد المستفيد

من المكتبة على الأقل ، إلى أي المصادر المرجعية يمكن أن يلجأ إليه للحصول على إجابة لسؤال معين. وأقرب مثال لهذه النظم أنسرمان (Answerman (Waters, 1986) الذي تم تصميمه في المكتبة الزراعية القومية . وقد روعي في هذا التصميم أن تعمل قوائم الاختيار على تضيق مجال سؤال المستفيد ، وأن تقوده نحو نوعية الأداة (الدليل أو معجم الأماكن ، أو المعجم المتخصص ، إلخ) اللازمة للإجابة عن السؤال. وينظر ووترز (Waters (1992 إلى الرد على الاستفسارات باعتباره تطبيقاً واضحاً لأساليب النظم الخبيرة ، نظراً لتكرار توجيه الأسئلة المتشابهة ، ونظراً لأن بعض المكتبات تسجل الأسئلة التي تتلقاها والإجابات التي تقدمها ، وبذلك يتكون « رصيد معرفي » ملائم. وقد أعد باروت (Parrott (1992 مراجعة علمية شاملة وتصنيفاً للنظم الخبيرة التي صممت للمساعدة في عملية الإرشاد في المكتبات.

ومنذ سنوات عدة مضت ، وعندما قدم لون (Luhn (1959 لأول مرة وصفاً لإحدى طرق تقديم خدمة البث الانتقائي للمعلومات باستخدام الحاسبات الآلية ، فقد وضع تصوراً لنظام يمكن أن يتعلم من أخطائه . فمن الممكن لسمات اهتمامات المشاركين في برنامج البث الانتقائي للمعلومات أن تتعدل آلياً استجابة لتقييمها للمواد المسترجعة . كما يمكن أيضاً الارتفاع بوزن المصطلحات الواردة في السمات ، أو خفض وزنها ، تبعاً لما إذا كانت ترتبط بالمواد التي أقر المتلقي صلاحيتها أو عدم صلاحيتها . كما يمكن أيضاً استبعاد أي مصطلح من السمات إذا تكرر رفض من يتلقى خدمة البث الانتقائي المواد المسترجعة بهذا المصطلح نظراً لعدم صلاحيتها . وكان من الصعب تنفيذ طريقة لون الآلية لتجديد السمات ، كما أنها لم تطبق كاملة في النظم العاملة على الإطلاق ، هذا على الرغم من أنه لا يوجد من حيث المبدأ ، سبب لعدم قابليتها للتطبيق. ولو قدر لها أن تطبق لكان من الممكن القول بأنها تبدي قدراً من الذكاء نظراً لأنها تتعلم من أخطائها.

ولا يتمتع بأي من خصائص التعلم الحقيقي سوى عدد ضئيل جداً من النظم التي تسمى « بالخبيرة » أو التي تنطوي على ذكاء اصطناعي ، في مجال المكتبات . والاستثناء الجدير بالتنويه في هذا السياق هو التطبيق الذي وصفه بونتيجو وآخرون (Pontigo et al. (1992 : فقد صمم هذا النظام الذي حظي بالوصف لمساعدة المكتبي على أن يحدد من أي مصدر يمكن طلب كتاب أو وعاء معين من أوعية المعلومات. ويقوم رصيد المعرفة المستخدم في

هذا النظام بربط بيانات التحقق من الوثائق (كالأرقام المعيارية الدولية للكتب وأرقام التقارير التقنية على سبيل المثال) بالمصادر المحتملة للتوريد . ويقال إن هذا النظام يتسم « بالدينامية والقابلية للتكيف » ، حيث يمكن إدخال البيانات الخاصة بمعدلات النجاح في الحصول على نوعيات معينة من أوعية المعلومات من موردين بعينهم، إلى النظام لتجديد رصيده المعرفي ، ومن ثم زيادة احتمالات الحصول على وعاء معين من المورد الذي يقع عليه الاختيار.

وقد لخص فنلي (1992) Fenly المزايا المحتملة للنظم الخبيرة والتي أمكن التحقق منها في الإنتاج الفكري؛ فالنظم الخبيرة :

- ١ - تتيح الخبرة النادرة على نطاق واسع ، ومن ثم تساعد غير الخبراء على تحقيق نتائج مناظرة لتلك التي يحققها الخبراء.
- ٢ - تكفل للخبراء فرصة استثمار جزء من وقتهم في أنشطة أخرى .
- ٣ - تدعم مقومات التقييس والاطراد في مهام تفتقر نسبياً إلى الانتظام .
- ٤ - توفر حوافز إنشاء مرصد بيانات معرفي ، في شكل دائم (لا يتوقف على سبيل المثال، على توافر أفراد بعينهم) .
- - تعمل على مستوى مرتفع وبشكل مطرد (لا تتأثر مثلاً بالإجهاد أو عدم القدرة على التركيز) .

وكل هذه المزايا حقيقية ولا مبالغة فيها ، وما من شك أنه يمكن للنظم الخبيرة التي يتم تصميمها بعناية ، أن تكون لها أهميتها بالنسبة للعمل بالمكتبات ، عندما تطبق في الأنشطة المفرقة في التخصص ، والتي لا تتم الآن إلا بإنفاق قدر كبير من الوقت من جانب خبراء من البشر مرتفعي التكلفة. كذلك يمكن للنظم الخبيرة أن تفيد في المهام التي يمكن أن تستثمر وبوضوح القدرة على التعلم ، كأنشطة طلب الوثائق مثلاً ، التي تناولها بونتيجو وآخرون (1992) Pontigo et al. إلا أنه ليس هناك ما يدعم الاعتقاد بأنه سيكون بإمكان الأجهزة المتمتعة « بالذكاء » أن تتولى وبسرعة كثيراً من المهام الفكرية التي يضطلع بها الآن المكتبيون المتمرسون المؤهلون تأهيلاً جيداً ، كما يبدي كثير من الكتاب في هذا الموضوع تفاؤلاً زائداً عن الحدود فيما يتعلق بهذه النقطة .

ويقول متسلر (1992) Metzler على سبيل المثال :

« ربما تكون مكتبة المستقبل قادرة على توفير مقومات الوصول إلى المعرفة التي تشتمل عليها مجموعاتنا (بشكل ضمنى في غالب الأحيان) والإفادة من هذه المعرفة بشكل أكثر ثراء وكثافة . وربما كانت نظم استرجاع المعلومات القائمة على المحتوى هي أبرز مظاهر التقدم في هذا الاتجاه . وقد يتطلب ذلك بالطبع ، نوعاً من الذكاء الاصطناعي ، وقدرة على فهم اللغة الطبيعية ، أكثر اتساعاً وأشد قوة مما هو متاح لنا الآن . ويمكن للخطوة التالية أن تنطوي لا على فهم النص بشكل جيد لتبين ما إذا كان صالحاً بالنسبة لحاجة عامة إلى المعلومات ، أعرب عنها المستفيد فحسب ، وإنما تنطوي أيضاً على فهم النص بشكل جيد يكفل القدرة على اقتطاف المعلومات التي يمكن أن تستخدم من جانب البرنامج » (8 - 9 PP) .

ويذكرنا التمس السائد للذكاء الاصطناعي ، في بعض قطاعات مجال المكتبات الآن ، بما ساد بعض قطاعات مهنة الطب منذ حوالي عشرين عاماً ، من تمس للتشخيص بمساعدة الآلات . ولم يحظ التشخيص بمساعدة الآلات بالقبول على نطاق واسع في الأوساط الطبية ، كما تبين الآن أن المشكلات أكثر صعوبة مما بدت وقتئذ . فالخبراء من البشر يمارسون عملهم بمزيج من المعرفة والخبرة والحدس ، واختزان المعرفة في أي شكل إلكتروني أمر ممكن إن لم يكن يسيراً ، أما تسجيل الخبرة البشرية فمشكلة أكثر صعوبة ، ولا يمكن أن يكون هناك في المستقبل القريب بديل للحدس البشري .

وتتطلب معظم الأنشطة التي يقوم بها المكتبيون ، قدرأ من المعرفة والخبرة والحدس أقل مما يتطلبه التشخيص الطبي ، ورغم ذلك فإن المشكلات التي تنطوي عليها أتمة أبسط المهام الفكرية قلما تحظى بالتقدير المناسب . فكما بين ديفز (1986) Davies وهو على حق : « فإن الخبرة في الفهرسة لا تتجلى في القواعد ، وإنما تكمن في الخطوات المتتابعة المترتبة على بعضها البعض والتي يتخذها الخبراء الذين يقومون بالعمل » (58 P) . كذلك نجد واييل (1992) Weibel ، وفي إشارة إلى العمل الذي قام به بوركو وإرسيجوفاك (1989) Borko and Ercegovac في فهرسة الخرائط ، ينوه بالنتائج العامة

أو الخلاصة التي خرجا بها ، وهي أن الخبرة اللازمة في مثل هذه الإجراءات « تتجاوز حدود ما هو منصوص عليه بوضوح في مجموعات القواعد المعتمدة » (P. 71) وأن تعقد النشاط يقف حجر عثرة في سبيل تطبيق أساليب النظم الخبرة.

وواقع الأمر أنه لا يمكن بالطبع ، تفويض الآلات بسهولة ، الاضطلاع بالمهام الفكرية الحقيقية المرتبطة بأعمال المكتبات ، والتي عرضنا لها تفصيلاً في هذا الكتاب، من تحليل موضوعي ، وتفسير للحاجة إلى المعلومات ، واستراتيجيات البحث ، وما شابه ذلك . وأياً كان ما يمكن أن يحدث للمكتبات كمؤسسات ، أي كمجموعات من المنتجات المادية، فإنه لا يمكن لخبرة المكتبيين المتمرسين أن يستعاض عنها بالذكاء الاصطناعي أو تقنية أخرى في المستقبل القريب . وكما عبر هورتون (Horton (1982) عن ذلك بفصاحة: « فإن الإبداع والموهبة ورجاحة العقل ... وليست آلات معالجة المعلومات ، هي (الأرصدة الرأسمالية) الحقيقية لإدارة المعلومات » (P. 39) .

الوراقية

- Alberico, Ralph, and Mary Micco. *Expert Systems for Reference and Information Retrieval*. Wesport, Conn., Meckler, 1990.
- Albright, John B. «Some Limits to Subject Retrieval from a Large Published Index». Doctoral dissertation. Urbana, Ill., University of Illinois, Graduate School of Library Science, 1979.
- Al-Hawamdeh S., et al. «Best Match Document Retrieval: Development and Use of INSTRUCT». In: *Online Information 88. 12th International Online Information Meeting, London, 6-8 December 1988. Proceedings*. Vol. 2. Oxford, Eng., Learned Information, 1988, pp. 761 - 777.
- Allen, Bryce L. «Bibliographic and Text-Linguistic Schemata in the User-Intermediary Interaction». Doctoral dissertation. London, Ont., Can., University of Western Ontario, School of Library and Information Science, 1988.
- Allen, Thomas J., and P. G. Gerstberger. *Criteria for Selection of an Information Source*. Cambridge, Mass., Massachusetts Institute of Technology, Sloan School of Management, 1966. Another version appears in *Journal of Applied Psychology*, 52 (4): 272 - 279, 1968.
- Alligood, Elaine C.; Elaine Russo-Martin; and Richard Peterson. «Use Study of Excerpta Medica Abstract Journals: To Drop or Not to Drop?» *Bulletin of the Medical Library Association*, 71(3): 251 - 258, July 1983.
- Anderson, Charles R. «Online Ready Reference in the Public Library». In: *Questions and Answers: Strategies for Using the Electronic Reference Collection. Proceedings of the 1987 Clinic on Library Applications of Data Processing*. Edited by Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1989, pp. 71-84.
- Anderson, James D. «Essential Decisions in Indexing Systems Design». In: *Indexing Specialized Formats and Subjects*. Edited by Hilda Feinberg. Metuchen, N.J., Scarecrow Press, 1983, pp. 1-21.
- Ashmole, R. F.; D. E. Smith; and B. T. Stern. «Cost Effectiveness of Current Awareness Sources in the Pharmaceutical Industry». *Journal of the American Society for Information Science*, 24 (1): 29 - 39, January/February 1973.

-
- Atkinson, Stephen D., and Michael Knee. *Subject Index to Databases Available from Computer Search Service*. Albany, N. Y., State University of New York, University Libraries, 1986. ED 267 825
- Austin, Derek. *PRECIS: A Manual of Concept Analysis and Subject Indexing*. 2nd ed. London, Eng., The British Library, 1984.
- Bannon, Cynthia. «The Perseus Project». In: *Hypertext/Hypermedia Handbook*. Edited by Emily Berk and Joseph Devlin. New York, McGraw-Hill, 1991, pp. 480-487.
- Bar, Jacob. «The Multifile Multidisciplinary (Horizontal) Search Approach—Justification and Principles». *Online Review*, 12 (1): 47-58, February 1988.
- Barber, John, et al. «Case Studies of the Indexing and Retrieval of Pharmacology Papers». *Information Processing and Management*, 24 (2): 141-150, 1988.
- Barron, Billy. *The Barron's Guide to Internet Accessible Library Catalogs*. Denton, Tex., University of North Texas, 1992 (available via anonymous FTP on the node FTP.unt.edu).
- Bates, Marcia J. «How to Use Information Search Tactics Online». *Online*, 11 (3): 47-54, May 1987.
- . «Subject Access in Online Catalogs: a Design Model». *Journal of the American Society for Information Science*, 37 (6): 357-376, November 1986.
- . «Idea Tactics». *Journal of the American Society for Information Science*, 30 (5): 280-289, September 1979 a.
- . «Information Search Tactics». *Journal of the American Society for Information Science*, 30 (4): 205 - 214, July 1979 b.
- . «System Meets User: Problems in Matching Subject Search Terms». *Information Processing and Management*, 13 (6): 367-375, 1977.
- Baxendale, Phyllis B. «Machine-Made Index for Technical Literature-An Experiment». *IBM Journal of Research and Development*, 2 (4): 354-361, October 1958.
- Belkin, Nicholas J. «Anomalous States of Knowledge as a Basis for Information Retrieval». *Canadian Journal of Information Science*, 5:133-143, May 1980.
- Belkin, Nicholas J., et al. «ASK for Information Retrieval: Part I. Background and Theory». *Journal of Documentation*, 38 (2): 61-71, June 1982 a.
- . «ASK for Information Retrieval: Part II. Results of a Design Study». *Journal of Documentation*, 38 (3): 145-164, September 1982 b.
- Bellardo, Trudi. «An Investigation of Online Searcher Traits and Their Relationships to Search Outcome». *Journal of the American Society for Information Science*, 36 (4): 241-250, July 1985.
- Belzer, Jack. «Information Theory as a Measure of Information Content». *Journal of the American Society for Information Science*, 24 (4): 300-304, July-August 1973.
- Berger, Mary C., and Judith Wanger. «Retrieval, Analysis and Display of Numeric Data». *Drexel Library Quarterly*, 18 (3/4): 11-26, Summer/Fall 1982.
- Berk, Emily, and Joseph Devlin. «What is Hypertext? In: *Hypertext/Hypermedia Handbook*. Edited by Emily Berk and Joseph Devlin. New York, McGraw Hill, 1991, pp. 3-7.
- Bernier, Charles L. «Correlative Indexes. 1. Alphabetical Correlative Indexes». *American Documentation*, 7 (4): 283-288, October 1956.
- Black, William J. «Knowledge-Based Abstracting». *Online Review*, 14(5): 327-340, October 1990.
-

- Blackshaw, Lyn, and Baruch Fischhoff. «Decision Making in Online Searching». *Journal of the American Society for Information Science*, 39 (6): 369-389, November 1988.
- Blair, David C. *Language and Representation in Information Retrieval*. Amsterdam, Neth., Elsevier Science, 1990.
- . «Full Text Retrieval: Evaluation and Implications». *International Classification*, 13 (1): 18-23, 1986.
- Blair, David C., and M. E. Maron. «An Evaluation of Retrieval Effectiveness for a Full-Text Document System». *Communications of the Association for Computing Machinery*, 28 (3): 289-299, March 1985.
- Borgman, Christine L. «Why Are Online Catalogs Hard to Use? Lessons Learned from Information-Retrieval Studies». *Journal of the American Society for Information Science*, 37 (6): 387-400, November 1986.
- . «The User's Mental Model of an Information Retrieval System: Effects on Performance». Doctoral dissertation. Palo Alto, Calif., Stanford University, 1984.
- Borgman, Christine L.; D. Moghdam; and P. K. Corbett. *Effective Online Searching: A Basic Text*. New York, Marcel Dekker, 1984.
- Borko, Harold, and M. Bemick. «Automatic Document Classification». *Journal of the Association for Computing Machinery*, 10: 151-162, 1963.
- Borko, Harold, and Zorana Ercegovac. «Knowledge-Based Descriptive Cataloging of Cartographic Publications». In: *Annual Review of OCLC Research, July 1988-June 1989*. Dublin, Ohio, OCLC, 1989, pp. 49-50.
- Bourne, Charles P. *Characteristics of Coverage by the Bibliography of Agriculture of the Literature Relating to Agricultural Research and Development*. Palo Alto, Calif., Information General Corporation, 1969a. PB 185425.
- . *Overlapping Coverage of the Bibliography of Agriculture by Fifteen Other Secondary Sources*. Palo Alto, Calif., Information General Corporation, 1969b. PB 185069.
- . «Some User Requirements Stated Quantitatively in Terms of the 90 Percent Library». In: *Electronic Information Handling*. Edited by Allen Kent and Orrin E. Taulbec. Washington, D. C., Spartan Books, 1965, pp. 93-110.
- Boyce, Bert R., and John P. McLain. «Entry Point Depth and Online Search Using a Controlled Vocabulary». *Journal of the American Society for Information Science*, 40 (4): 273-276, July 1989.
- Bradford, S. C. *Documentation*. 2nd ed. London, Eng., Crosby Lockwood, 1953.
- British Standards Institution. *British Standard Guide to Establishment and Development of Monolingual Thesauri*. London, Eng., British Standards Institution, 1987. Bs 5723
- . *Guidelines for the Establishment and Development of Monolingual Thesauri*. London, Eng., British Standards Institution, 1979.
- Brittain, J. Michael, and S. A. Roberts. «Rationalization of Secondary Services: Measurement of Coverage of Primary Journals and Overlap Between Services». *Journal of the American Society for Information Science*, 31 (3): 131-142, May 1980.
- Britten, William A. «A Use Statistics for Collection Management: The 80/20 Rule Revisited». *Library Acquisitions: Practice and Theory*, 14 (2): 183-189, 1990.
- Brooks, H. M. «Expert Systems and Intelligent Information Retrieval». *Information Processing and Management*, 23 (4): 367-382, 1987.
- Brown, Steven Allan. «Telefacsimile in Libraries: New Deal in the 1980's». *Library Trends*, 37 (3): 343-356, Winter 1989.

- Bryant, Edward C.; Donald W. King; and P. James Terragno. *Some Technical Notes on Coding Errors*. Bethesda, Md., Westat Research, Inc., 1963. PB 166487.
- BSO Referral Index: a Subject Index to 36 Data-bases in the DIALOG System*. The Hague, Neth., Fédération Internationale de Documentation, 1985.
- Buchanan, Lori E.; Anne May Berwind; and Don Carlin. «Optical Disk-Based Periodical Indexes for Undergraduates». *College and Research Libraries News*, 50 (1): 10-14, January 1989.
- Burgin, Robert. «The Effect of Indexing Exhaustivity on Retrieval Performance». *Information Processing and Management*, 27 (6): 623-628, 1991.
- Bush, Vannevar. «As We May Think» *The Atlantic Monthly*, 176 (1): 101-108, July 1945.
- Busha, Charles H., and Stephen P. Harter. *Research Methods in Librarianship: Techniques and Interpretation*. New York, Academic Press, 1980.
- Byler, Anne Meyer, and Mary Ravenhall. «Using Dialindex for the Identification of On-Line Databases Relevant to Urban and Regional Planning». *Online Review*, 12 (2): 119-133, April 1988.
- Calkins, Mary L. «Free Text or Controlled Vocabulary? A Case History Step-by-Step Analysis...Plus Other Aspects of Search Strategy». *Database*, 3 (2): 53-67, June 1980.
- Campbell, Donald T., and Julian C. Stanley. *Experimental and Quasi-Experimental Design for Research*. Chicago, Ill., Rand-McNally, 1963.
- Carrow, Deborah, and Joan Nugent. «Comparison of Free-Text and Index Search Abilities in an Operating Information System». In: *Proceedings of the American Society for Information Science, 40th Annual Meeting, Chicago, Ill., September 26 - October 1, 1977*. Vol. 14. White Plains, N. Y., Knowledge Industry Publications, 1977. (On microfiche card 2 of 10, E8-F3).
- Cawkell, A. E. «Selected Aspects of Image Processing and Management: Review and Future Prospects». *Journal of Information Science*, 18 (3): 179-192, 1992.
- Chamis, Alice Y. «Selection of Online Databases Using Switching Vocabularies». *Journal of the American Society for Information Science*, 39 (3): 217-218, May 1988.
- Cleverdon, Cyril W. «Optimizing Convenient Online Access to Bibliographic Databases». *Information Services and Use*, 4 (1 - 2): 37-47, April 1984.
- . *Report on the Testing and Analysis of an Investigation into the Comparative Efficiency of Indexing Systems*. Cranfield, Eng., College of Aeronautics, 1962.
- Cleverdon, Cyril W.; Jack Mills; and E. Michael Keen. *Factors Determining the Performance of Indexing Systems*. 3 vols. Cranfield, Eng., College of Aeronautics, 1966.
- Cochrane, Pauline A. *Improving LCSH for Use in Online Catalogs*. Littleton, Colo., Libraries Unlimited, 1986.
- Collette, A. D., and J. A. Price. «A Cost/Benefit Evaluation of Online Interactive Bibliographic Searching in a Research and Engineering Organization». In: *The Value of Information: Collection of Papers Presented at the Sixth Mid-Year Meeting, American Society for Information Science, Syracuse University, Syracuse, N. Y., May 19-21, 1977*. Washington, D. C., American Society for Information Science, 1977, pp. 24-34.
- Cooper, Marianne. «Current Information Dissemination: Ideas and Practices». *Journal of Chemical Documentation*, 8 (4): 207-218, November 1968.

- Cooper, William S. «On Selecting a Measure of Retrieval Effectiveness». *Journal of the American Society for Information Science*, 24 (2): 87-100, March-April 1973.
- . «A Definition of Relevance for Information Retrieval». *Information Storage and Retrieval*, 7 (1): 19-37, June 1971.
- . «Is Interindexer Consistency a Hobgoblin?» *American Documentation*, 20 (3): 268-278, July 1969.
- Courrier, Yves. «SYNTOL». In: *Encyclopedia of Libray and Information Science*. Vol. 29. Edited by Allen Kent et al. New York, Marcel Dekker, 1980, pp. 357-381.
- Corft, W. Bruce; Howard R. Turtle; and David D. Lewis. «The Use of Phrases and Structured Queries in Information Retrieval». In: *Proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval, 13-16 October, 1991, Chicago, Illinois*. Edited by A. Bookstein, Y. Chiaramella, G. Salton, and V. V. Raghavan. New York, Association for Computing Machinery, 1991, pp. 32-45.
- Crum, Norman J. «The Librarian-Customer Relationship: Dynamics of Filling Requests for Information». *Special Libraries*, 60 (5): 269-277, May-June 1969.
- Cuadra Associates. *Directory of Online Databases*. Vol. 13, No. 2. Edited by K. Y. Marcaccio. Detroit, Mich., Cuadra/Gale, 1992.
- Cuadra, Carlos A., and Robert V. Katter. «Opening the Black Box of 'Relevance'». *Journal of Documentation*, 23 (4): 291-303, December 1967a.
- . «The Relevance of Relevance Assessment». In: *Proceedings of the American Documentation Institute*. Vol. 4. Washington, D. C., Thompson Book Company, 1967n, pp. 95-99.
- Cuadra, Carlos A.; Robert V. Katter; Emory H. Holmes; and Everett M. Wallace. *Experimental Studies of Relevance Judgments: Final Report*. 3 vols. Santa Monica, Calif., System Development Corporation, 1967.
- Cutter, Charles A. *Rules for a Dictionary Catalog*. Washington, D.C., U.S. Government Printing Office, 1876.
- Dana, John Cotton. *Libraries: Addresses and Essays*. White Plains, N.Y., H. W. Wilson Co., 1916.
- Das Gupta, Padmini. «Boolean Interpretations of Conjunctions for Document Retrieval». *Journal of the American Society for Information Science*, 38 (4): 245-254, July 1987.
- Davies, Roy. «Expert Systems and Cataloguing: New Wine in Old Bottles?» In *Expert Systems in Libraries*. Edited by Forbes Gibb. London, Eng., Taylor Graham, 1986, pp. 67-82.
- Davison, P. S., and D. A. R. Matthews. «Assessment of Information Services». *Aslib Proceedings*, 21 (7): 280-283, July 1969.
- Dialog Information Services, Inc. *Search Strategy Seminar*. Palo Alto, Calif., Dialog Information Services, Inc., 1983.
- Diaz, Lily «PathMAC: An Alternative Approach to Medical Education». In: *Hypertext/Hypermedia Handbook*. Edited by Emily Berk and Joseph Devlin. New York, McGraw Hill, 1991, pp. 488-492.
- Dillon, Martin and Ann S. Gray. «FASIT: A Fully Automatic Syntactically Based Indexing System». *Journal of the American Society for Information Science*, 34 (2): 99-108, March 1983.
- Dillon, Martin, and Laura McDonald. «Fully Automatic Book Indexing». *Journal of Documentation*, 39 (3): 135-145, September 1983.

- Doszkoć, Tamas E. «Natural Language Processing in Information Retrieval». *Journal of the American Society for Information Science*, 37 (4): 191-196, July 1986.
- . «CITE NLM: Natural-Language Searching in an Online Catalog». *Information Technology and Libraries*, 2 (4): 364-380, December 1983.
- . «AID, an Associative Interactive Dictionary for Online Searching». *Online Review*, 2(2): 163-173, June 1978.
- Dougherty, Richard M., and Carol Hughes. *Preferred Futures for Libraries: A Summary of Six Workshops with University Provosts and Library Directors*. Mountain View, Calif., Research Libraries Group, 1991.
- Dowlin, Kenneth E. «The Neographic Library». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F. W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Doyle, Lauren B. «Semantic Road Maps for Literature Searchers». *Journal of the Association for Computing Machinery*, 8 (4): 553-578, October 1961.
- Dykstra, Mary. «Can Subject Headings Be Saved?» *Library Journal*, 113 (15): 55-58, September 15, 1988a.
- . «LC Subject Headings Disguised as a Thesaurus». *Library Journal*, 113 (4): 42-46, March 1, 1988b.
- Dym, Eleanor D. «Relevance Predictability: I. Investigation Background and Procedures». In: *Electronic Handling of Information: Testing and Evaluation*. Edited by Allen Kent et al. Washington, D.C., Thompson Book Co., 1967, pp. 175-185.
- Edmundson, H.P. «New Methods in Automatic Extracting». *Journal of the Association for Computing Machinery*, 16 (2): 264-285, April 1969.
- Eisenberg, Michael, and Carol Berry. «Order Effects: a Study of the Possible Influence of Presentation Order on User Judgments of Document Relevance». *Journal of the American Society for Information Science*, 39 (5): 293-300, September 1988.
- Elchesen, Dennis R. «Cost-Effectiveness Comparison of Manual and On-Line Retrospective Bibliographic Searching» *Journal of the American Society for Information Science*, 29 (2): 56-66, March 1978.
- El-Shooky, E., et al. *Selecting the Most Appropriate Databases to Answer Industrial Information Requests*. The Hague, Neth., International Federation for Information and Documentation, 1988.
- Estabrook, Leigh Stewart. «Valuing a Document Delivery System». *RQ*, 26 (1): 58-62, Fall 1986.
- Evans, David A., et al. «Automatic Indexing Using Selective NLP and First-Order Thesauri». In: *Proceedings of RIAO 1991-Intelligent Text and Image Handling, 2-5 April 1991, Barcelona, Spain*. Edited by A. Lichnerowicz. Amsterdam, Neth., Elsevier, 1991, pp. 624-643.
- Fagan, Joel L. *Experiments in Automatic Phrase Indexing for Document Retrieval: A Comparison of Syntactic and Non-Syntactic Methods*. Ithaca, N.Y., Cornell University, 1988.
- Fairthorne, R.A. Unpublished notes. 1965.
- Fenichel, Carol Hansen. «Online Searching: Measures that Discriminate Among Users With Different Types of Experiences». *Journal of the American Society for Information Science*, 32 (1): 23-32, January 1981.
- . «An Examination of the Relationship Between Searching Behavior and Searcher Background» *Online Review*, 4 (4): 341-347, December 1980a.

- . «The Process of Searching Online Bibliographic Databases: a Review of Research». *Library Research*, 2 (2): 107-127, Summer 1980b.
- Fenly, Charles. «Technical Services Processes as Models for Assessing Expert System Suitability and Benefits». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 50-66.
- Fidel, Raya. «Who Needs Controlled Vocabulary?» *Special Libraries*, 83 (1): 1-9, Winter 1992.
- . «Searchers' Selection of Search Keys». *Journal of the American Society for Information Science*, 42 (7): 490-527, August 1991.
- . «Writing Abstracts for Free-Text Searching». *Journal of Documentation*, 42 (1): 11-12, March 1968.
- . «Individual Variability in Online Search Behavior». In: *Proceedings of the 48th American Society for Information Science Annual Meeting, Las Vegas, Nevada, October 20-24, 1985*. Vol. 22. Edited by Carol A. Parkhurst. White Plains, N.Y., Knowledge Industry Publications, 1985, pp. 69-72.
- . «Online Searching Styles: a Case-Study-Based Model of Searching Behavior». *Journal of the American Society for Information Science*, 35 (4): 211-221, July 1984.
- Field, B. J. «Towards Automatic Indexing: Automatic Assignment of Controlled-Language Indexing and Classification from Free Indexing». *Journal of Documentation*, 31(4): 246-265, December 1975.
- Fitzgerald, Evelyn L. C. «The Value of the Search Request Form in the Negotiation Process Between Requester and Librarian». Doctoral dissertation. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1981.
- Flowerdew, A. D. J., and C. M. E. Whitehead. *Cost-Effectiveness and Cost Benefit Analysis in Information Science*. London, Eng., London School of Economics and Political Science, 1974. OSTI Report No. 5206.
- Foskett, D. J. «A Note on the Concept of 'Relevance.'» *Information Storage and Retrieval*, 8 (2): 77-78, April 1972.
- . «Classification and Indexing in the Social Sciences». *Aslib Proceedings*, 22 (3): 90-101, March 1970.
- Fox, Edward A.; Bernard Rous; and Gary Marchionini. «ACM's Hypertext and Hypermedia Publishing Projects». In: *Hypertext/Hypermedia Handbook*. Edited by Emily Berk and Joseph Devlin. New York, McGraw Hill, 1991, pp. 460-464.
- Fox, Edward A., et al. «Building a Large Thesaurus for Information Retrieval». In: *Proceedings of the Second Conference on Applied Natural Language Processing, Austin, Texas, February 9-12, 1988*. Morristown, N. J., Association for Computational Linguistics, 1988, pp. 101-108.
- Fussler, Herman H., and Julian L. Simon. *Patterns in the Use of Books in Large Research Libraries*. 2nd ed. Chicago, Ill., University of Chicago Press, 1969.
- Gardin, Jean Claude. *SYNTOL*. New Brunswick, N.J., Rutgers, the State University, 1965.
- Gauch, Susan, and John B. Smith. «An Expert System for Automatic Query Reformation». *Journal of the American Society for Information Science*, 44(3): 124-136, 1993.
- Geller, Valerie, and Michael E. Lesk. «An On-Line Library Catalog Offering Menu and Keyword User Interfaces». In: *Proceedings of the Fourth National Online Meeting*,

-
- New York, April 12-14, 1983. Medford, NJ., Learned Information, 1983, pp. 159-165.
- Goffman, William. «On Relevance as a Measure». *Information Storage and Retrieval*, 2(3): 201-203, December 1964.
- Goffman, William, and V. A. Newill. *Methodology for Test and Evaluation of Information Retrieval Systems*. Cleveland, Ohio, Case Western Reserve University, Center for Documentation and Communication Research, 1964.
- Goldhor, Herbert. *An Introduction to Scientific Research in Librarianship*. Urbana, Ill., University of Illinois, Graduate School of Library Science, 1972. Graduate School of Library Science Monograph No. 12.
- Graham, Deborah L. «Simultaneous Remote Search». *Bulletin of the Medical Library Association*, 68 (4): 370-371, October 1980.
- Greenberg, Bette, et al. «Evaluation of a Clinical Medical Librarian Program at the Yale Medical Library». *Bulletin of the Medical Library Association*, 66 (3): 319-326, July 1978.
- Hanwen, Kathleen A. «The Effect of Presearch Experience on the Success of Naive (End-User) Searches». *Journal of the American Society for Information Science*, 37 (5): 315-318, September 1986.
- Harmon, Donna. «The DARRA TIPSTER Project». *SIGIR Forum*, 26 (2): 26-28, 1992.
- Harter, Stephen P. «Detrimental Effects of Searching with Precoordinated Terms». *Online Review*, 12 (4): 205-210, August 1988.
- . «Online Information Retrieval: Concepts, Principles and Techniques». Orlando, Fla., Academic Press, 1986.
- . «Online Searching Styles: an Exploratory Study». *College and Research Libraries*, 45 (4): 249-258, July 1984.
- Havener, W. Michael. «Answering Ready Reference Questions: Print Versus Online». *Online*, 14 (1): 22-28, January 1990.
- Hawkins, Donald T. «Applications of Artificial Intelligence (AI) and Expert Systems for Online Searching». *Online*, 12 (1): 31-43, January 1988.
- Heaps, Doreen M., and Paul Sorenson. «An On-Line Personal Documentation System». In: *Proceedings of the American Society for Information Science, 31st Annual Meeting, Columbus, Ohio, October 20-24, 1968*. Vol. 5. New York, Greenwood Publishing, 1968, pp. 201-207.
- Heller, Martin. «Future Documents». *Byte*, 16 (5): 126-129, 132, 134-135, May 1991.
- Hirschman, Lynette; Ralph Grishman; and Naomi Sager. «Grammatically - Based Automatic Word Class Formation». *Information Processing and Management*, 11(1/2): 39-57, June 1975.
- Hitch, Charles J., and Roland McKean. *The Economics of Defense for the Nuclear Age*. Cambridge, Mass., Harvard University, 1960.
- Holm, B. E., and L. E. Rasmussen. «Development of a Technical Thesaurus». *American Documentation*, 12 (3): 184-190, July 1961.
- Horton, Forest W., Jr. «Human Capital Investment: Key to the Information Age». *Information and Records Management*, 16 (7): 38-39, July 1982.
- Howard, Helen. «Measures that Discriminate Among Online Searchers with Different Training and Experience». *Online Review*, 6 (4): 315-327, August 1982.
- Hsieh-Yee, Ingrid. «The Search Tactics of Novice and Experienced Searchers». Doctoral dissertation. Madison, Wis., University of Wisconsin, Madison, 1990.
-

- Hu, Chengren. «An Evaluation of a Gateway System for Automated Online Database Selection». In: *Proceedings of the Ninth National Online Meeting, New York, N.Y., May 10-12, 1988*. Medford, N.J., Learned Information, 1988, pp. 107-114.
- . «An Evaluation of Online Database Selection by a Gateway System with Artificial Intelligence Techniques». Doctoral dissertation. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1987.
- Huang, Samuate T., and Terrence J. McHale. «A Cost-Effectiveness Comparison Between Print and Online Versions of the Same Frequently-Used Sources of Business and Financial Information». In: *Proceedings of the Eleventh National Online Meeting, New York, N.Y., May 1-3, 1990*. Edited by Martha E. Williams. Medford, N.J., Learned Information, 1990, pp. 161-168.
- Hulme, E. Wyndham. «Principles of Book Classification: An Introduction». *Library Association Record*, 13:354-356, 1911a.
- . «Principles of Book Classification: Chapter II. Principles in Book Classification». *Library Association Record*, 13:389-394, 1911b.
- . «Principles of Book Classification: Chapter III. On the Definition of Class Headings, and the Natural Limit to the Extension of Book Classification». *Library Association Record*, 13: 444-449, 1911c.
- Humphrey, Susanne M. «Interactive Knowledge-Based Systems for Improved Subject Analysis and Retrieval». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 81-117.
- Humphrey, Susanne M., and Nancy E. Miller. «Knowledge-Based Indexing of the Medical Literature: The Indexing Aid Project». *Journal of the American Society for Information Science*, 38(3): 184-196, May 1987.
- Hutchins, W. J. «The Concept of 'Aboutness' in Subject Indexing». *Aslib Proceedings*, 30 (5): 172-181, May 1978.
- Ide, E., and Gerard Salton. «Interactive Search Strategies and Dynamic File Organization». In: *The SMART Retrieval System - Experiments in Automatic Document Processing*. Edited by Gerard Salton. Englewood Cliffs, N.J., Prentice Hall, 1971, pp. 373-393.
- Ingwersen, Peter. «Search Procedures in the Library-Analysed from the Cognitive Point of View». *Journal of Documentation*, 38(3): 165-191, September 1982.
- International Organization for Standardization. *Guidelines for the Establishment and Development of Monolingual Thesauri*. Geneva, Switzerland, ISO, 1986. ISO 2788.
- Janes, Joseph W. «Relevance Judgments and the Incremental Presentation of Document Representations». *Information Processing and Management*, 27(6): 629-646, 1991.
- Jeng, Ling-Hwey. «An Expert System for Determining Title Proper in Descriptive Cataloging: A Conceptual Model». *Cataloging and Classification Quarterly*, 7(2):55-70, Winter 1986.
- Jensen, Rebecca J.; Herbert O. Asbury; and Radford G. King. «Costs and Benefits to Industry of Online Literature Searches». *Special Libraries*, 71(7): 291-299, July 1980.
- Jones, Kevin P. «Problems Associated with the Use of Compound Words in Thesauri With Special Reference to BS 5723: 1979». *Journal of Documentation*, 37(2): 35-68, June 1981.

-
- . «Compound Words: A Problem in Post Coordinate Retrieval Systems». *Journal of the American Society for Information Science*, 22 (4): 242-250, July/August 1971.
- Katz, William. «The Reference Interview and Levels of Service». In: *Introduction to Reference Work*. Vol. 2. New York, McGraw Hill, 1987, pp. 39-58.
- Katzer, Jeffery, et al. «A Study of the Overlap Among Document Representations». *Information Technology: Research and Development*, 1 (4): 261-274, October 1982.
- Keen, E. Michael. «Evaluation Parameters». In: *The SMART Retrieval System: Experiments in Automatic Document Processing*. Edited by Gerard Salton. Englewood Cliffs, N. J., Prentice Hall, 1971, pp. 74-111.
- . *Measures and Averaging Methods Used in Performance Testing of Indexing Systems*. Cranfield, Eng., Aslib Cranfield Research Project, 1966.
- Kemp, D. A. «Relevance, Pertinence and Information System Development». *Information Storage and Retrieval*, 10(2): 37-49, February 1974.
- Kent, Allen, et al. «Relevance Predictability in Information Retrieval Systems». *Methods of Information in Medicine*, 6 (2): 45-51, April 1967.
- . «Machine Literature Searching. VIII. Operational Criteria for Designing Information Retrieval Systems». *American Documentation*, 6(2): 39-101, April 1955.
- Kilgour, Frederick G. «The Metamorphosis of Libraries During the Foreseeable Future». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F. W. Lancaster. Binghamton, N. Y., Haworth Press, 1993.
- King, David N. «The Contribution of Hospital Library Information Services to Clinical Care: a Study in Eight Hospitals.» *Bulletin of the Medical Library Association*, 75 (4): 291-301, October 1987.
- King, Donald W. *Comments on the Meaning and Interpretation of Consistency Measures for Evaluating Indexing Processes*. Bethesda, Md., Westat Research, Inc., 1967.
- . «Evaluation of Coordinate Index Systems During File Development». *Journal of Chemical Documentation*, 5 (2): 96-99, May 1965.
- King, Donald W., and Edward C. Bryant. *The Evaluation of Information Services and Products*. Arlington, Va., Information Resources Press, 1971.
- King, Donald W., et al. *The Value of the Energy Data Base*. Rockville, Md., King Research, Inc., 1982. DOE/OR / 11232 - 1 - (DE 82014250).
- King, Geraldine B. «The Reference Interview». *RQ*, 12 (2): 157-160, Winter 1972.
- Klingbiel, Paul H. «A Technique for Machine-Aided Indexing». *Information Storage and Retrieval*, 9 (9): 477-494, September 1973.
- Knapp, Sara D. «The Reference Interview in the Computer-Based Setting». *RQ*, 17 (4): 320-324, Summer 1978.
- Kochen, M. «Organizing Knowledge for Coping with Needs». In: *Ordering Systems for Global Information Networks: Proceedings of the Third International Study Conference on Classification Research, Bombay, India, 6-11 January 1975*. Edited by A. Neelameghan. Bangalore, India, FID/CR and Sarada Ranganathan Endowment for Library Science, 1979, pp. 142-149.
- Konings, C. A. G. «Comparison and Evaluation of Nine Bibliographies/Bibliographic Databases in the Field of Computer Science». *Online Review*, 9 (2): 121-133, April 1985.
- Kramer, Joseph. «How to Survive in Industry: Cost Justifying Library Services». *Special Libraries*, 62 (11): 487-489, November 1971.
-

- Kremer, Jeannette M. «Perspectives for Information Services and Professionals in Brazil». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F.W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Krol, E. *The Whole Internet: User's Guide and Catalog*. Sebastopol, Calif., O'Reilly and Associates, 1992.
- Kurzweil, Raymond. «Machine Intelligence: The First 80 years». *Library Journal*, 116 (13): 69-71, August 1991.
- Lancaster, F. W. *If You Want to Evaluate Your Library ...* 2nd ed. Urbana Ill., University of Illinois, Graduate School of Library and Information Science, 1993a.
- . ed. *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Binghamton, N. Y., Haworth Press, 1993b.
- . *Indexing and Abstracting in Theory and Practice*. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1991.
- . *Vocabulary Control for Information Retrieval*. 2nd. Arlington, Va., Information Resources Press, 1986.
- . «The Paperless Society Revisited». *American Libraries*, 16(8): 553-555, September 1985.
- . *Libraries and Librarians in an Age of Electronics*. Arlington, Va., Information Resources Press, 1982.
- . «Some Considerations Relating to the Cost-Effectiveness of Online Services in Libraries». *Aslib Proceedings*, 33 (1): 10-14, January 1981.
- . *Toward Paperless Information Systems*. New York, Academic Press, 1987.
- . «The Information Services Librarian». *Australian Special Libraries News*, 7 (6): 139-149, November 1974a.
- . «A Study of Current Awareness Publications in the Neurosciences». *Journal of Documentation*, 30 (3): 255-272, September 1974b.
- . *Vocabulary Control for Information Retrieval*. Arlington, Va., Information Resources Press, 1972.
- . «Aftermath of an Evaluation». *Journal of Documentation*, 27 (1): 1-10, March 1971.
- . *Evaluation of the MEDLARS Demand Search Service*. Bethesda, Md., National Library of Medicine, 1968a.
- . Letter to the Editor. *American Documentation*, 19 (2): 206, April 1968b.
- Lancaster, F. W., and Ja-Lih Lee. «Bibliometric Techniques Applied to Issues Management: A Case Study». *Journal of the American Society for Information Science*, 36 (6): 389-397, November 1985.
- Lancaster, F. W., and Jack Mills. «Testing Indexes and Index Language Devices: The Aslib Cranfield Project». *American Documentation*, 15 (1): 4-13, January 1964.
- Lancaster, F. W., et al. «Searching Databases on CD-ROM: Comparison of the Results of End User Searching with Results from Two Modes of Searching by Skilled Intermediaries». «Submitted for publication in RQ, 1993.
- . *Searching Databases on CD-ROM: Comparison of the Results of End User Searching With Results from Two Modes of Searching by Skilled Intermediaries. A report to the Council on Library Resources*. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992.
- . «Identifying Barriers to Effective Subject Access in Library Catalogs». *Library Resources and Technical Services*, 35 (4): 377-391, October 1991.

- Landau, Herbert B. «International Ownership in the U.S. Information Marketplace: Myths, Paradoxes, and Ironies». *Information Services and Use*, 8 (2/3/4): 36-71, 1988.
- Larson, Ray R. «Evaluation of Advanced Retrieval Techniques in an Experimental Online Catalog». *Journal of the American Society for Information Science*, 43 (1): 34-53, 1992.
- Ledwith, Robert. «On the Difficulties of Applying the Results of Information Retrieval to Aid in the Searching of Large Scientific Databases». *Information Processing and Management*, 28 (4): 451-455, July-August 1992.
- Lefever, Maureen; Barbara Freedman; and Louise Schultz. «Managing and Uncontrolled Vocabulary Ex Post Facto». *Journal of the American Society for Information Science*, 23 (6): 339-342, November/December 1972.
- Leonard, Lawrence E. «Inter-Indexer Consistency and Retrieval Effectiveness: Measurement of Relationships». Doctoral dissertation. Urbana, Ill., University of Illinois, Graduate School of Library Science, 1975.
- Lesk, Michael E. «Word-Word Associations in Document Retrieval Systems». *American Documentation*, 20 (1): 27-38, 1969.
- Lesk, Michael E., and Gerard Salton. «Interactive Search and Retrieval Methods Using Automatic Information Displays». In: *AFIPS Conference Proceedings, 1969 Spring Joint Computer Conference, Boston, Mass., May 14-16, 1969*. Montvale, N.J., AFIPS Press, 1969, pp. 345-446.
- . «Relevance Assessments and Retrieval System Evaluation». *Information Storage and Retrieval*, 4 (4): 343-359, December 1968.
- Lewis, D. D., and W. B. Croft. «Term Clustering of Syntactic Phrases». In: *Proceedings of the 13th International Conference on Research and Development in Information Retrieval, 5-7 September 1990, Brussels, Belgium*. Edited by Jean-Luc Vidick. New York, Association for Computing Machinery, 1990, pp. 385-404.
- Lewis, David W. «Inventing the Electronic University». *College and Research Libraries*, 49 (4): 291-304, July 1988.
- Lilley, Oliver L. «Evaluation of the Subject Catalog». *American Documentation*, 5(2):41-60, April 1954.
- Line, Maurice B. «Libraries and Information Services in 25 Years' Time: A British Perspective». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F.W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- . «The Ability of a University Library to Provide Books Wanted by Researchers». *Journal of Librarianship*, 5 (1): 37-51, January 1973.
- Line, Maurice B., and A. Sandison. «Obsolescence' and Changes in the Use of Literature With Time». *Journal of Documentation*, 30 (3): 283-350, September 1974.
- Littleford, Alan. «Artificial Intelligence and Hypermedia». In: *Hypertext/Hypermedia Handbook*. Edited by Emily Berk and Joseph Devlin. New York, McGraw Hill, 1991, pp. 357-378.
- Logan, Elisabeth L. «Cognitive Styles and Online Behavior of Novice Searchers». *Information Processing and Management*, 26 (4): 503-510, 1990.
- Lottor, M. *Internet Growth (1981-1991)*. Palo Alto, Calif., Network Information Systems Center, SRI International, 1992.
- Lucky, Robert W. *Silicon Dreams: Information, Man, and Machine*. New York, St. Martin's Press, 1989.

-
- Luhn, Hans Peter. *Selective Dissemination of New Scientific Information with the Aid of Electronic Processing Equipment*. Yorktown Heights, N.Y., IBM Advanced Systems Development Division, 1959.
- . «The Automatic Creation of Literature Abstracts». *IBM Journal of Research and Development*, 2(2): 159-165, April 1958.
- . «A Statistical Approach to Mechanized Encoding and Searching of Literary Information». *IBM Journal of Research and Development*, 1(4): 309-317, October 1957.
- Lynch, Mary Jo. «Reference Interviews in Public Libraries», *Library Quarterly*, 48(2): 119-142, April 1978.
- Magson, M. S. «Techniques for the Measurement of Cost-Benefit in Information Centers». *Aslib Proceedings*, 25 (5): 164-185, May 1973.
- Mandel, Carol A. «Trade-offs: Quantifying Quality in Library Technical Services». *Journal of Academic Librarianship*, 14 (4): 214-220, September 1988.
- Mandersloot, Wim G. B.; Eleanor M. B. Douglal; and Neville Spicer. «Thesaurus Control - The Selection, Grouping and Cross-Referencing of Terms for Inclusion in a Coordinate Index Word List». *Journal of the American Society for Information Science*, 21(1): 49-57, January/February 1970.
- Marcaccio, Kathleen Young; Juli Adams; and Martha E. Williams, eds. *Computer-Readable Databases: A Directory and Data Sourcebook*. 6th ed. Detroit, Mich., Gale Research, 1990.
- Marchetti, Pier Giorgio, and Nicholas J. Belkin. «Interactive Online Search Formulation Support». In: *Proceedings of the Twelfth National Online Meeting, New York, N.Y., May 7-9, 1991*. Edited by Martha E. Williams. Medford, N.Y., Learned Information, 1991, pp. 237-243.
- Marcus, Richard S., and Fred J. Reintjes. *Computer Interface for User Access to Heterogeneous Information Retrieval Systems*. Cambridge, Mass., MIT Electronic Systems Laboratory, 1977. (ESL-R-739).
- Marcus, Richard S.; Alan R. Benenfeld; and Peter Kugel. «The User Interface for the Intrex Retrieval System». In: *Interactive Bibliographic Search: The User/Computer Interface*. Edited by Donald E. Walker. Montvale, N.J., AFIPS Press, 1971, pp.159-201.
- Markey, Karen. «Interindexer Consistency Tests: A Literature Review and Report of a Test of Consistency in Indexing Visual Materials». *Library and Information Science Research*, 6(2): 155-177, April/June 1984a.
- . *Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs*. Dublin, Ohio, OCLC, Inc., 1984b.
- . «Levels of Question Formulation in Negotiation of Information Need During the Online Presearch Interview: a Proposed Model». *Information Processing and Management*, 17(5): 215-225, 1981.
- Markey, Karen, and Pauline Cochrane. *ONTAP: Online Training and Practice Manual for ERIC Data Base Searchers*. 2nd ed. Syracuse, N.Y., Syracuse University, ERIC Clearinghouse on Information Resources, 1981.
- Markey, Karen; Pauline Atherton; and Claudia Newton. «An Analysis of Controlled Vocabulary and Free Text Search Statements in Online Searches». *Online Review*, 4(3): 225-236, September 1980.

-
- Maron, M. E. «On Indexing, Retrieval and the Meaning of About». *Journal of the American Society for Information Science*, 28(1): 38-43, January 1977.
- Marron, Harvey. «On Costing Information Services». In: *Proceedings of the American Society for Information Science, 32nd Annual Meeting, San Francisco, Calif., October 1-4, 1969*. Vol. 6. Edited by Jeanne B. North. Westport, Conn., Greenwood Publishing, 1969, pp. 515-520.
- Martyn, John «Tests on Abstracts Journals: Coverage, Overlap, and Indexing». *Journal of Documentation*, 23(1): 45-70, March 1967.
- . «Unintentional Duplication of Research». *New Scientist*, 21(377): 338, February 6, 1964.
- Martyn, John, and F.W. Lancaster. *Investigative Methods in Library and Information Science*. Arlington, Va., Information Resources Press, 1981.
- Martyn, John, and Margaret Slater. «Tests on Abstracts Journals». *Journal of Documentation*, 20 (4): 212-235, December 1964.
- Martyn, John; Peter Vickers; and Mary Feeney, eds. *Information UK 2000*. London, Eng., Bowker-Saur, 1990.
- Marx, Gray T., and Sanford Sherizen. «Ethical Issues: Monitoring on the Job.» In: *Computers in the Human Context*. Edited by Tom Forester. Cambridge, Mass., MIT Press, 1989, pp. 397-406.
- Mason, Donald. «PPBS: Application to an Industrial Information and Library Service». *Journal of Librarianship*, 4 (2): 91-105, April 1972.
- Mathis, Betty A.; James E. Rush; and Carol E. Young. «Improvement of Automatic Abstracts by the Use of Structural Analysis». *Journal of the American Society for Information Science*, 24 (2): 101-109, March/April 1973.
- Matthews, Joseph R.; Gary S. Lawrence; and Douglas K. Ferguson. *Using Online Catalogs: A Nationwide Survey*. New York, Neal-Schuman, 1983.
- McCarthy, Martin V. «InfoMaster: a Powerful Information Retrieval Service for Business». *Online*, 10 (6): 53-58, November 1986.
- McCue, Janice H. *Online Searching in Public Libraries: a Comparative Study of Performance*. Metuchen, N.J., Scarecrow Press, 1988.
- McKinin, Emma Jean, et al. «The Medline/Full-Text Research Project». *Journal of the American Society for Information Science*, 42 (4): 297-307, May 1991.
- Meadows, Donella H., et al. *The Limits to Growth*. New York, Universal Books, 1972.
- Metzler, Douglas P. «Artificial Intelligence: What Will They Think of Next?» In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F.W. Lancaster and Linda C. Smith Urbana, Ill. University of Illinois, Graduate School of Library and Information Science, 1992, pp. 2-49.
- Meyer, Daniel E., and Den Ruiz. «End-User Selection of Databases - Part I: Science/Technology/Medicine». *Database*, 13 (3): 21-29, June 1990a.
- . «End-User Selection of Databases - Part II: Business/Law». *Database*, 13 (4): 35-42, August 1990b.
- . «End-User Selection of Databases - Part III: Social Science/Arts & Humanities». *Database*, 13 (5): 59-64, October 1990c.
- Milstead, Jessica. *Subject Access Systems: Alternatives in Design*. Orlando, Fla., Academic Press, 1984.

- Mitchell, Maurice, and Laverne M. Saunders «The Virtual Library: An Agenda for the 1990s». *Computers in Libraries*, 11 (4): 8, 10-11, April 1991.
- Molholt, Patricia. «Libraries as Bridges: Librarians as Builders». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F.W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Mondschein, Lawrence G. «SDI Use and Productivity in the Corporate Research Environment». *Special Libraries*, 81 (4): 265-279, Fall 1990.
- Morris, A.; G. Tseng; and G. Newham. «The Selection of Online Databases and Hosts-an Expert System Approach». In: *Online Information 88: 12th International Online Information Meeting, London, 6-8 December 1988. Proceedings*. Vol. 1. Oxford, Eng., Learned Information, 1988, pp. 139-148.
- Morrison, Margaret. «The NISO Common Command Language». *Online*, 13 (4): 46-52, July 1989.
- Mueller, M.W. «Time, Cost and Value Factors in Information Retrieval». Paper Presented at the IBM Information Systems Conference, Poughkeepsie, N.Y., September 21-23, 1959.
- Mylonas, Elli. «An Interface to Classical Greek Civilization». *Journal of the American Society for Information Science*, 43 (2): 192-201, 1992.
- Neelameghan, A. «Libraries and Information Services in Third World Countries». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F.W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Nelson, Theodor H. «Getting It Out of Our System». In: *Information Retrieval: A Critical View*. Edited by George Schechter. Washington, D.C., Thompson Book Co., 1967, pp. 191-210.
- Nicholas, David; Gertrud Erbach; and Kevin Harris. «Online: Views on Costs and Cost-Effectiveness». *Journal of Information Science*, 13 (2): 109-115, 1987.
- Nicholls, Paul, and Ria Van Den Elshout. «Survey of Databases Available on CD-ROM: Types, Availability, and Content». *Database*, 13 (1): 18-23, February 1990.
- Niehoff, Robert, and Greg Mack. «The Vocabulary Switching System. Description of Evaluation Studies». *International Classification*, 12 (1): 2-6, 1985.
- Nielsen, Jakob. «The Art of Navigating Through Hypertext». *Communications of the Association for Computing Machinery*, 33 (3): 296-310, March 1990.
- Nightingale, R. A. «A Cost-Benefit Study of a Manually-Produced Current Awareness Bulletin». *Aslib Proceedings*, 25 (4): 153-157, April 1973.
- Nixon, Judith M. «Online Searching of Human Nutrition: An Evaluation of Databases». *Medical Reference Services Quarterly*, 8 (3): 27-35, Fall 1989.
- O'Brien, Ann. «Relevance as an Aid to Evaluation in OPACs». *Journal of Information Science*, 16 (4): 265-271, 1990.
- OCLC, Inc. *Assessing Information on the Internet: Toward Providing Library Services for Computer-Mediated Communication*. Proposal to the U.S. Department of Education. May 1991 - January 1992.
- . *OCLC Annual Report: 1990-1991*. Dublin Ohio, OCLC Online Computer Library Center, 1991.
- O'Connor, John. «Some Independent Agreements and Resolved Disagreements About Answer-Providing Documents». *American Documentation*, 20 (4): 311-319, October 1969.

-
- . «Some Questions Concerning 'Information Need'». *American Documentation*, 19 (2): 200-203, April 1968a.
- . Letter to the Editor. *American Documentation* 19 (4): 416-417, October 1968b.
- . «Relevance Disagreements and Unclear Request Forms». *American Documentation*, 18 (3): 165-177, July 1967.
- O'Leary, Mick. «Easynet Revisited: Pushing the Online Frontier». *Online*, 12(5): 22-30, September 1988.
- . «DIALOG Business Connection: DIALOG for the End-User». *Online*, 10 (5): 15-24, September 1986.
- Oliver, Lawrence H., et al. *An Investigation of the Basic Processes Involved in the Manual Indexing of Scientific Documents*. Bethesda, Md., General Electric Co.. Information Systems Operation, 1966. PB 169415.
- Paice, Chris D. «Constructing Literature Abstracts by Computer: Techniques and Prospects». *Information Processing and Management*, 26(1): 171-186, 1990.
- Pao, Miranda L. *Concepts of Information Retrieval*. Englewood, Colo., Libraries Unlimited, 1989.
- Parker, Lorraine M. P., and Robert E. Johnson. «Does Order of Presentation Affect Users' Judgment of Documents?» *Journal of the American Society for Information Science*, 41 (7): 493-494, October 1990.
- Parrott, James R. «Reference Expert Systems: Foundations in Reference Theory». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 188-160.
- Penhale, Sara J., and Nancy Taylor. «Integrating End-User Searching into a Bibliographic Instruction Program». *RQ*, 26 (2): 212-220, Winter 1986.
- Penniman, W. David. «Shaping the Future for Libraries through Leadership and Research». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F. W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Pollitt, Steven. «CANSEARCH: An Expert System Approach to Document Retrieval». *Information Processing and Management*, 23 (2): 119-138, 1987.
- Pontigo, Jaime, et al. «Expert Systems in Document Delivery: The Feasibility of Learning Capabilities». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 254-266.
- Potter, William Gray. «Expanding the Online Catalog». *Information Technology and Libraries*, 8 (2): 99-104, June 1989.
- Powell, Ronald R. *Basic Research Methods for Librarians*. Norwood, N.J., Ablex, 1985.
- Price, Derek J. de Solla. *Little Science, Big Science*. New York, Columbia University Press, 1963.
- . *Science Since Babylon*. New Haven, Conn., Yale University Press, 1961.
- Quade, E. S. *Systems Analysis Techniques for Planning-Programming-Budgeting*. Santa Monica, Calif, The Rand Corporation, 1966. P-3322.
-

- Quarterman, John S. *The Matrix: Computer Networks and Conferencing Systems Worldwide*. Bedford, Mass., Digital Press, 1990.
- Rada, Roy; Hafedh Mill; Gary Letourneau; and Doug Johnston. «Creating and Evaluating Entry Terms». *Journal of Documentation*, 44 (1): 19-41, March 1988.
- Reffel, Jeffery S., and Robert Shishko. *Systematic Analysis of University Libraries*. Cambridge, Mass., MIT Press, 1969.
- Raitt, David. «The Library of the Future». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F.W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Rath, G. J., et al. «Comparison of Four Types of Lexical Indicators of Content». *American Documentation*, 12 (2): 126-130, April 1961.
- Rees, Alan M. «The Relevance of Relevance to the Testing and Evaluation of Document Retrieval Systems». *Aslib Proceedings*, 18 (11): 316-324, November 1966.
- Rees, Alan M., and Tesko Saracevic. «The Measurability of Relevance». *Proceedings of the American Documentation Institute. Annual Meeting, Santa Monica, Calif, October 3-7, 1966*, Vol. 3. Woodland Hills, Calif, Adrienne Press, 1966, pp. 225-234.
- Rees, Alan M., and Douglas G. Schultz. *A Field Experimental Approach to the Study of Relevance Assessments in Relation to Document Searching: Final Report*. 2 Vols. Cleveland, Ohio, Case Western Reserve University, 1967.
- Reich, Phyllis, and Erik J. Biever. «Indexing Consistency: The Input/Output Function of Thesauri». *College and Research Libraries*, 52 (4): 336-342, July 1991.
- Resnick, A. «Relative Effectiveness of Document Titles and Abstracts for Determining Relevance of Documents». *Science*, 134 (3484): 1004-1006, October 6, 1961.
- Robertson, S. E. «The Parametric Description of Retrieval Tests». *Journal of Documentation*, 25 (2): 93-107, June 1969.
- Rogers, Michael. «First Electronic Medical Journal to Debut in 1992». *Library Journal*, 116 (18): 32, 34, November 1, 1991.
- Roloff, Michael E. «Communication at the User-System Interface: a Review of Research». *Library Research*, 1 (1): 1-18, Spring 1979.
- Rosenberg, Kenyon C. «Evaluation of an Industrial Library: A Simple-Minded Technique». *Special Libraries*, 60 (10): 635-638, December 1969.
- Rosenberg, Victor. «The Application of Psychometric Techniques to Determine the Attitudes of Individuals Toward Information Seeking». Bethlehem, Pa., Lehigh University, Center for Information Sciences, 1966. Another version appears in *Information Storage and Retrieval*, 3 (3): 119-127, July 1967.
- Rowlett, Russell J., Jr. «Keywords vs. Index Terms». *Journal of Chemical Information and Computer Sciences*, 17 (3): 192-193, August 1977.
- Ruge, Gerda. «Experiments on Linguistically-Based Term Associations». *Information Processing and Management*, 28 (3): 317-332, 1992.
- Ruiz, Den, and Daniel E. Meyer. «End-User Selection of Databases - Part IV: People/ News/General Reference». *Database*, 13 (5): 65-67, October 1990.
- Rush, James E.; R. Salvador; and A. Zamora. «Automatic Abstracting and Indexing. II. Production of Indicative Abstracts by Application of Contextual Inference and Syntactic Coherence Criteria». *Journal of the American Society for Information Science*, 22 (4): 260-274, July/August 1971.
- Salton, Gerard. «The State of Retrieval System Evaluation». *Information Processing and Management*, 28 (4): 441-449, 1992.

-
- . *Automatic Text Processing: The Transformation, Analysis and Retrieval of Information by Computer*. Reading, Mass., Addison-Wesley, 1989.
- . «Syntactic Approaches to Automatic Book Indexing». In: *Proceedings of the 2nd Applied Natural Language Processing Conference, 9-12 February 1988, Austin, Texas*. Morristown, N.J., Association for Computational Linguistics, 1988, pp. 204-210.
- . ed, *The SMART Retrieval System: Experiments in Automatic Document Processing*. Englewood Cliffs, N.J., Prentice-Hall, 1971.
- . *Automatic Information Organization and Retrieval*. New York, McGraw-Hill, 1968.
- Salton, Gerard, and Michael J. McGill. *Introduction to Modern Information Retrieval*. New York, McGraw Hill, 1983.
- Salton, Gerard; C. Buckley; and E. A. Fox. «Automatic Query Formulations in Information Retrieval». *Journal of the American Society for Information Science*, 34 (4): 262-280, July 1983.
- Sandison, A. «The Use of Older Literature and Its Obsolescence». *Journal of Documentation*, 27 (3): 184-199, September 1971.
- Saracevic, Tefko. «Individual Differences in Organizing, Searching and Retrieving Information». In: *Proceedings of the American Society for Information Science, 54th Annual Meeting, Washington, D. C., October 27-31, 1991*. Vol. 28. Edited by José-Marie Griffiths. Medford, N.J., Learned Information, 1991, pp. 82-86.
- . «Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science». *Journal of the American Society for Information Science*, 26 (6): 321-343, November / December 1975.
- . «On the Concept of Relevance in Information Science». Doctoral Dissertation. Cleveland, Ohio, Case Western Reserve University, 1970a.
- . «Ten Years of Relevance Experimentation - A Summary and Synthesis of Conclusions». In: *Proceedings of the American Society for Information Science, 33rd Annual Meeting, Philadelphia, Pa., October 11-15, 1970*. Vol. 7. Edited by Jeanne B. North. Washington, D. C., American Society for Information Science, 1970b, pp. 33-36.
- . «Comparative Effects of Titles, Abstracts and Full Texts on Relevance Judgments». In: *Proceedings of the American Society for Information Science, 32nd Annual Meeting, Philadelphia, Pa., October 1-4, 1969*. Vol. 6. Edited by Jeanne B. North. Westport, Conn., Greenwood Publishing, 1969, pp. 293-299.
- Saracevic, Tefko, and Paul Kantor. «A Study of Information Seeking and Retrieving: II. Users, Questions, and Effectiveness». *Journal of the American Society for Information Science*, 39 (3): 177-196, May 1988a.
- . «A Study of Information Seeking and Retrieving: III. Searchers, Searches, and Overlap». *Journal of the American Society for Information Science*, 39 (3): 197-216, May 1988b.
- Saracevic, Tefko, et al. «A Study of Information Seeking and Retrieving: I. Background and Methodology». *Journal of the American Society for Information Science*, 39 (3): 161-176, May 1988.
- Schamber, Linda; Michael B. Eisenberg; and Michael S. Nilan. «A Re-Examination of Relevance: Toward a Dynamic, Situational Definition». *Information Processing and Management*, 26 (6): 755-776, 1990.
-

- Schnall, Janet G., and Joan W. Wilson. «Evaluation of a Clinical Medical Librarianship Program at a University Health Sciences Library». *Bulletin of the Medical Library Association*, 64 (3): 278-283, July 1976.
- Schrage, Michael. *Shared Minds: The New Technologies of Collaboration*. New York, Random House, 1990.
- Schwarz, Helga. «Expert Systems and the Future of Cataloguing: a Possible Approach». *LIBER Bulletin*, 26 : 23-50, 1986.
- Scura, Georgia, and Frank Davidoff. «Case-Related Use of the Medical Literature: Clinical Librarian Services for Improving Patient Care». *Journal of the American Medical Association*, 245 (1): 50-52, January 2, 1981.
- Seiler, Lauren H., and Thomas T. Surprenant. «The Virtual Information Center: Scholars and Information in the 21st Century». In: *Libraries and the Future: Essays on the Library in the Twenty-first Century*. Edited by F. W. Lancaster. Binghamton, N.Y., Haworth Press, 1993.
- Self, Phyllis C.; Thomas W. Filardo; and F. W. Lancaster. «Acquired Immunodeficiency Syndrome (AIDS) and the Epidemic Growth of Its Literature». *Scientometrics*, 17 (1/2): 49-60, July 1989.
- Sharma, V. S. «A Comparative Evaluation of Online Databases in Relation to Welfare and Corrective Services, and Community Development». *Online Review*, 6 (4): 297-313, August 1982.
- Shaw, W. M., Jr. «Subject Indexing and Citation Indexing. Part I: Clustering Structure in the Cystic Fibrosis Document Collection». *Information Processing and Management*, 26 (6): 693-703, 1990a.
- . «Subject Indexing and Citation Indexing. Part II: An Evaluation and Comparison». *Information Processing and Management*, 26 (6): 705-718, 1990b.
- Shirey, Donald L., and Marvin Kurfeerst. «Relevance Predictability: II. Data Reduction». In: *Electronic Handling of Information: Testing and Evaluation*. Edited by Allen Kent et al. Washington, D.C., Thompson Book Company, 1967, pp. 187-198.
- Shurkin, Joel. *Engines of the Mind: A History of the Computer*. New York, W.W. Norton 1984.
- Sievert, MaryEllen, and Mark J. Andrews. «Indexing Consistency in Information Science Abstracts». *Journal of the American Society for Information Science*, 42 (1): 1-6, January 1991.
- Sievert, MaryEllen, and Bert R. Boyce. «Hedge Trimming and the Resurrection of the Controlled Vocabulary in Online Searching». *Online Review*, 7(6):489-494, December 1983.
- Smith, Linda C. «Artificial Intelligence in Information Retrieval». In: *Annual Review of Information Science and Technology*. Edited by Martha E. Williams. Vol. 22. Amsterdam, Neth., Elsevier, 1987, pp. 41-77.
- . «Knowledge-Based Systems, Artificial Intelligence and Human Factors». In: *Information Technology and Information Use: Towards a Unified View of Information and Information Technology, May 8-10, 1985, Copenhagen, Denmark*. Edited by P. Ingwersen, L. Kajberg, and A. M. Pejtersen. London, Eng., Taylor Graham, 1986, pp. 98-110.
- Smith, Linda C., and Amy J. Warner «A Taxonomy of Representations in Information Retrieval System Design». *Journal of Information Science*, 8 (3): 113-121, April 1984.

- Smith, Sallye Wry. «Venn Diagramming for On-Line Searching». *Special Libraries*, 67 (11): 510-517, November 1976.
- Snow, Bonnie. «Database Selection in the Life Sciences». *Database*, 8 (3): 15-44, August 1985.
- Soergel, Dagobert. *Indexing Languages and Thesauri: Construction and Maintenance*. Los Angeles, Calif., Melville Publishing Co., 1974.
- Solomin, V.M. «Efficiency Indexes for the Performance of Information Agencies». English translation appears in *Scientific and Technical Information Processing*, 2 : 16-23, 1974.
- Somerville, Arleen N. «The Place of the Reference Interview in Computer Searching: the Academic Setting». *Online*, 1 (4): 14-23, October 1977.
- Sommar, Helen G., and Don E. Dennis. «A New Method of Weighted Term Searching with a Highly Structured Thesaurus». In: *Proceedings of the American Society for Information Science, 32nd Annual Meeting, San Francisco, Calif., October 1-4, 1969*. Vol. 6. Edited by Jeanne B. North Westport, Conn., Greenwood Publishing, 1969, pp. 193-198.
- Sparck Jones, Karen, ed. *Information Retrieval Experiment*. London, Eng., Butterworth, 1981.
- . *Automatic Keyword Classification for Information Retrieval*. London, Eng., Butterworth, 1971.
- . «Experiments in Semantic Classification». *MT: Mechanical Translation*, 8 (3/4): 97-112, June/October 1955.
- Stallings, Warren D; Sarah E. Hutchinson; and Stacey C. Sawyer. *Computers: The User Perspective*. St. Louis, Mo., Times Mirror/Mosby, 1988.
- Starr, Susan S. «Databases in the Marine Sciences». *Online Review*, 6(2): 109-126, April 1982.
- Stiles, H. Edmund. «The Association Factor in Information Retrieval». *Journal of the Association for Computing Machinery*, 8 (2): 271-279, April 1961.
- Sullivan, Michael V.; Christine L. Borgman; and Dorothy Wippem. «End-Users, Mediated Searches, and Front-End Assistance Programs on Dialog: A Comparison of Learning, Performance, and Satisfaction» *Journal of the American Society for Information Science*, 41 (1): 27-42, January 1990.
- Svenonius, Elaine. «Unanswered Questions in the Design of Controlled Vocabularies». *Journal of the American Society for Information Science*, 37 (5): 331-340, September 1986.
- Swanson, Don R. «Subjective Versus Objective Relevance in Bibliographic Retrieval Systems». *Library Quarterly*, 56 (4): 389-398, October 1986.
- . «Information Retrieval as a Trial-and-Error Process». *Library Quarterly*, 47 (2): 128-148, April 1977.
- Swets, John A. «Information Retrieval Systems». *Science*, 141(3577): 245-250, July 19, 1963.
- Swift, D. F.; V. Winn; and D. Bramer. «'Aboutness' as a Strategy for Retrieval in the Social Sciences». *Aslib Proceedings*, 30 (5): 182-187, May 1978.
- Tait, J. I., and Karen Sparck Jones. *Automatic Search Term Variant Generation for Document Retrieval*. Cambridge, Eng., University of Cambridge, Computer Laboratory, 1983. British Library R & D Report 5793.
- Taylor, Robert S. *Question Negotiation and Information-Seeking in Libraries*. Bethlehem,

- Pa., Lehigh University, Center for the Information Sciences, 1967. Also in *College and Research Libraries*, 29 (3): 178-194, May 1968.
- Tenopir, Carol. «Four Options for End User Searching». *Library Journal*, 111 (12): 56-57, July 1986.
- . «Full Text Database Retrieval Performance». *Online Review*, 9 (2): 149-164, April 1985.
- . «Retrieval Performance in a Full Text Journal Article Database». Doctoral Dissertation. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1984.
- . «Evaluation of Database Coverage: A Comparison of Two Methodologies». *Online Review*, 6 (5): 423-441, October 1982.
- Tenopir, Carol, and Gerald W. Lundeen. «Software Choices for In-House Databases». *Database*, 11 (3): 34-42, June 1988.
- Tenopir, Carol, and Jung Soon Ro. *Full Text Databases*. New York, Greenwood Press, 1990.
- Tesler, Lawrence G. «Networked Computing in the 1990s». *Scientific American*, 265 (3): 86-93, September 1991.
- Thompson, N.J. «DIALOGLINK and TRADEMARKSCAN - FEDERAL Pioneers in Online Images». *Online*, 13 (3): 15-26, May 1989.
- Thorpe, Peter. «An Evaluation of Index Medicus in Rheumatology: Coverage, Currency, and Efficiency». *Methods of Information in Medicine*, 13 (1): 44-47, January 1974.
- Trautman, Rodes, and Sara von Flittner. «An Expert System for Microcomputers to Aid Selection of Online Databases». *Reference Librarian*, 23 : 207-238, August 1989.
- Trueswell, Richard L. «Some Behavioral Patterns of Library Users: The 80/20 Rule». *Wilson Library Bulletin*, 43 (5): 458-461, January 1969.
- Ullmann, Stephen. *The Principles of Semantics*. 2nd ed. Oxford, Eng., Blackwell, 1963.
- Unesco. *Unesco Guidelines for the Establishment and Development of Monolingual Thesauri*. 2nd ed. (Prepared by Derek Austin and Peter Dale). Paris, 1981.
- van Brakel, P. A. «Easy Net: Intelligent Gateway to Online Searching». *South African Journal of Library and Information Science*, 56 (4): 191-197, September 1988a.
- . «Evaluating an Intelligent Gateway: a Methodology». *South African Journal of Library and Information Science*, 56 (4): 277-290, December 1988b.
- Van Rijsbergen, C. J. *Information Retrieval*. 2nd ed. London, Eng., Butterworth, 1979.
- Vickery, Alina. «The Experience of Building Expert Search Systems». In: *Online Information 88: 12th International Online Information Meeting, London, 6-8 December 1988. Proceedings*. Vol. 1. Oxford, Eng., Learned Information, 1988, pp. 301-313.
- Vickery, Alina, and H. M. Brooks. «PLEXUS - The Expert System for Referral». *Information Processing and Management*, 23 (2): 99-117, 1987.
- Vickery, Brian C. «Intelligent Interfaces to Online Databases». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F.W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 239-253.
- Visschedijk, Ankie, and Forbes Gibb. «Unconventional Text Retrieval Systems». *Online & CD-ROM Review*, 17 (1): 11-23, 1993.
- Vizine-Goetz, Diane, and Karen Markey Drabenstott. «Computer and Manual Analysis

- of Subject Terms Entered by Online Catalog Users». In: *Proceedings of the American Society for Information Science, 54th Annual Meeting, Washington, D.C., October 27-31, 1991*. Vol. 28. Edited by José-Marie Griffiths. Medford, N. J., Learned Information, 1991, pp. 156-161.
- Walker, Geraldene. «The Search Performance of End-Users». In: *Proceedings of the Ninth National Online Meeting, New York, N.Y., May 10-12, 1988*. Medford, N.J., Learned Information, 1988, pp. 304 - 410.
- Walker, J., and S. D. Atkinson. «Online Searching in the Humanities: Implications for End-Users and Intermediaries». In: *Online Information 1988, 12th International Online Information Meeting, London, 6-8 December 1988. Proceedings*. Vol. 1. Oxford, Eng., Learned Information, 1988, pp. 401-412.
- Walker, Stephen. «Improving Subject Access Painlessly: Recent Work on the OKAPI Online Catalogue Projects». *Program*, 22 (1): 21-31, January 1988.
- . «OKAPI: Evaluating and Enhancing an Experimental Online Catalog». *Library Trends*, 35 (4): 631-645, Spring 1987.
- Wang, Yih-Chen; James Vandendorpe; and Martha Evens. «Relational Thesauri in Information Retrieval». *Journal of the American Society for Information Science*, 36 (1): 15-27, January 1985.
- Wanger, Judith. «Multiple Database Use: the Challenge of the Database Selection Process». *Online*, 1 (4): 35-41, October 1977.
- Wanger, Judith; Dennis McDonald; and Mary C. Berger. *Evaluation of the On-Line Search Process*. Bethesda, Md., National Library of Medicine, 1980.
- Wamer, Amy J. «Natural Language Processing: Current Status for Libraries». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F.W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 194-214.
- . «A Linguistic Approach to Automatic Hierarchical Organization of Phrases». In: *Information in the Year 2000: From Research to Application, Proceedings of the 53rd Annual Meeting of the American Society for Information Science, 4-8 November, 1990, Toronto, Canada*. Edited by Diane Henderson. Medford, N.J., Learned Information, 1990, pp. 225-227.
- Waters, Samuel T. «Expert Systems at the National Agricultural Library: Past, Present, and Future» In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith Urbana, Ill., University of Illionis, Graduate School of Library and Information Science, 1992, pp. 161-177.
- . «Answerman, the Expert Information Specialish: An Expert System for Retrieval of Information from Library Reference Books». *Information Technology and Libraries*, 5 (3): 204-212, September 1986.
- Weber, Robert «The Clouded Future of Electronic Publishing». *Publishers Weekly*, 237 (26): 76-80, June 29, 1990.
- Weibel, Stuart. «Automated Cataloging: Implications for Libraries and Patrons». In: *Artificial Intelligence and Expert Systems: Will They Change the Library? Proceedings of the 27th Annual Clinic on Library Applications of Data Processing*. Edited by F. W. Lancaster and Linda C. Smith. Urbana, Ill., University of Illinois, Graduate School of Library and Information Science, 1992, pp. 67-80.

- Welsh, John J. «Evaluation of CD-ROM Use in a Government Research Library». *Laserdisk Professional*, 2 (6): 55-61, November 1989.
- Whitcomb, Laurie. «OCLC's EPIC System Offers a New Way to Search the OCLC Database». *Online*, 14 (1): 45-50, January 1990.
- White, Marilyn Domas. «Evaluation of the Reference Interview». *RQ*, 25 (1): 67-84, Fall 1985.
- . «The Reference Encounter Model». *Drexel Library Quarterly*, 19 (2): 38-55, Spring 1983.
- . «The Dimensions of the Reference Interview». *RQ*, 20 (4): 373-381, Summer 1981.
- Wildemuth, Barbara M., et al. «A Detailed Analysis of End-User Search Behaviors». In: *Proceedings of the American Society for Information Science, 54th Annual Meeting, Washington, D.C., October 27-31, 1991*. Vol. 28. Edited by José-Marie Griffiths. Medford, N.J., Learned Information, 1991, pp. 302-312.
- Williams, Martha E. «The State of Databases Today: 1992». In: *Computer-Readable Databases: Directory and Data Sourcebook*. 8th ed. Edited by Kathleen Young Marcaccio, Kevin Hillstrom, and Gwen E. Turecki. Detroit, Mich., Gale Research, 1992, pp. xi-xxi.
- . «Highlights of the Online Database Industry and the Quality of Information and Data». In: *Proceedings of the Eleventh National Online Meeting, New York, N. Y., May 1-3, 1990*. Edited by Martha E. Williams. Medford, N.J., Learned Information, 1990, pp. 1-4.
- . «Highlights of the Online Database Field: New Technologies for Online». In: *Proceedings of the Ninth National Online Meeting, New York, N. Y., May 10-12, 1988*. Edited by Martha E. Williams. Medford, N.J., Learned Information, 1988, pp. 1-4.
- . «Transparent Information Systems Thorough Gateways, Front Ends, Intermediaries and Interfaces». *Journal of the American Society for Information Science*, 37 (4): 204-214, July 1986.
- . «Electronic Databases». *Science*, 228 (4698): 445-456, April 26, 1985.
- . «Database and Online Statistics for 1979». *Bulletin of the American Society for Information Science*, 7 (2): 27-29, December 1980a.
- . «Future Directions for Machine-Readable Data Bases and Their Use». In: *Proceedings of the 1979 Clinic on Library Applications of Data Processing*. Edited by F.W. Lancaster. Urbana, Ill., University of Illinois, Graduate School of Library Science, 1980b, pp. 82-93.
- . *An Integrated Man/Machine Interface (Transparent System) to Facilitate Network Resource Utilization - a Feasibility Study*. Urbana, Ill., University of Illinois, Coordinated Science Laboratory, 1977.
- Wilson, Patrick. «The end of Specificity». *Library Resources and Technical Services*, 23 (2): 116-122, Spring 1979.
- . «Situational Relevance». *Information Storage and Retrieval*, 9 (8): 457-471, August 1973.
- Wilson, Sandra R.; Norma Starr-Schneidkraut; and Michael D. Cooper. *Use of the Critical Incident Technique to Evaluate the Impact of MEDLINE. Final Report*. Palo Alto, Calif., American Institute for Research, 1989. PB 90-142522
- Yonker, Valeria A., et al. «Coverage and overlaps in Bibliographic Databases Relevant to

-
- Forensic Medicine: A Comparative Analysis of MEDLINE». *Bulletin of the Medical Library Association*, 78 (1): 49-56, January 1990.
- Young, Philip H. «Visions of Academic Libraries in a Brave, New Future». In: *Libraries and the Future: Essays on the Library in the Twenty-First Century*. Edited by F. W. Lancaster. Binghamton, N. Y., Haworth Press, 1993.
- Zipf, George Kingsley. *The Psycho-Biology of Language*. Boston, Mass., Houghton-Mifflin, 1935.
- Zunde, Pranas, and Margaret E. Dexter. «Indexing Consistency and Quality». *American Documentation*, 20 (3): 259-267, July 1969.

قائمة المصطلحات

(A)

| | |
|-------------------------|-------------------|
| Aboutness | الحيثيات |
| Abstracting | الاستخلاص |
| Abstracts | المستخلصات |
| Acceptance rate | معدل القبول |
| Accessibility | سهولة المنال |
| Access points | المدخل |
| Accuracy | الدقة |
| Acronym | اسم استهلاكي |
| Affixes | اللواحق |
| Algorithm | خوارزمية |
| Ambiguity | الغموض |
| Approaches | أساليب التناول |
| Artificial intelligence | الذكاء الاصطناعي |
| Assignment indexing | التكشيف بالتعيين |
| Assimilation | الاستيعاب |
| Authority control | التقنين الاستنادي |
| Authorship | التأليف |

(B)

| | |
|------------------------|----------------------|
| Background information | المعلومات الأساسية |
| Batch processing | التجهيز على دفعات |
| Benefit analysis | تحليل العائد |
| Bibliographic item | وثيقة / وعاء معلومات |
| Bibliographic utility | مرفق وراقي |
| Bibliographic warrant | المسوغ الوراقى |
| Bibliometrics | القياسات الوراقية |
| Boolean logic | المنطق البوليني |

(C)

| | |
|----------------------|-----------------------------|
| CD - ROM | الأسطوانات الضوئية المكتنزة |
| Cataloguing | الفهرسة |
| Circulation | الإعارة |
| Classification | التصنيف |
| Clustering | التجميع العنقودي |
| Command language | لغة الأوامر |
| Comparative recall | الاستدعاء المقارن |
| Component | عنصر |
| Comprehensive search | البحث الشامل |
| Computer | الحاسب الآلى |
| Conceptual analysis | تحليل المفاهيم |
| Consistency | الاطراد |
| Content analysis | تحليل المضمون |
| Content words | كلمات المضمون |
| Coordination level | مستوى الربط |
| Copyright | حقوق التأليف |
| Cost benefit | عائد التكلفة |

Cost effectiveness
Current awareness
Cutoff

فعالية التكلفة
الإحاطة الجارية
قطاع

(D)

Database
Delegated search
Derivative indexing
Descriptors
Diminishing returns
Discrimination
Discrimination value
Disk
Dissemination
Document delivery
Document supply

مرصد البيانات
البحث المفوض
التكشيف الاشتقاقي
الواصفات
العائدات المتناقصة
التمييز
القيمة التمييزية
أسطوانة
البث
الإمداد بالوثائق
الإمداد بالوثائق

(E)

Electronic publishing
Electro-optic
End user
Entry terms
Entry vocabulary
Enumerative
Evaluation
Exhaustivity
Expert systems
Explode capability
Exponential growth

النشر الإلكتروني
الضوئية الإلكترونية
المستفيد النهائي
مصطلحات الدخول
لغة الدخول
حصري
التقييم
الشمول / الإحاطة
النظم الخبيرة
القدرة على التوسع
النمو الأسّي

| | |
|---------------------|-------------------|
| Extract | مقتطف |
| Extracting | الاقتطاف |
| Extraction indexing | التكشيف بالاقتطاف |

(F)

| | |
|---------------------|----------------------------|
| Fallout | السقط |
| False coordination | الربط الزائف |
| False drop | القطرة الزائفة |
| Feedback | التقييم المرتد |
| Formal channels | القنوات الرسمية |
| Fractional search | البحث الجزئي |
| Free-text searching | البحث في النصوص المطلقة |
| Full-text database | مرصد بيانات النصوص الكاملة |
| Function words | الكلمات الوظيفية |

(H)

| | |
|------------|--------------------------|
| Hardware | العتاد |
| Heuristic | إيعازي |
| Hit rate | معدل الإصابة |
| Homography | الجناس / الاشتراك اللفظي |
| Hub | نقطة ارتكازية |
| Hypertext | النصوص الفائقة |

(I)

| | |
|----------------|-------------------|
| Identifier | محدد |
| Indexing | التكشيف |
| Index language | لغة التكشيف |
| Index terms | المصطلحات الكشفية |

| | |
|----------------------------|----------------------------|
| Informal channels | القنوات غير الرسمية |
| Information broker | وسيط (سمسار) المعلومات |
| Information needs | الحاجة إلى المعلومات |
| Information retrieval | استرجاع المعلومات |
| Information services | خدمات / مرافق المعلومات |
| Information specialist | اختصاصي المعلومات |
| Information transfer cycle | دورة تداول المعلومات |
| Input | المدخلات |
| Intermediary | الوسيط |
| Interactive | تفاعلي |
| Interface | واجهة التعامل |
| Interlibrary loan | تبادل الإعارة بين المكتبات |
| Inverted file | الملف المصنف |
| Inverted index | الكشاف المصنف |
| Iterative searching | البحث التكراري |

(K)

| | |
|-----------------|----------------------------|
| Keywords | الكلمات المفتاحية / الدالة |
| Knowledge-based | قائم على المعرفة |

(L)

| | |
|----------------------|---------------------------|
| Least common factor | العامل النادر |
| Linear file | الملف التتابعي / التسلسلي |
| Linguistic groupings | التجميعات اللغوية |
| Links | الروابط |
| Literary warrant | المسوغ الأدبي |
| Literature | الإنتاج / النتاج الفكري |
| Literature searching | البحث في الإنتاج الفكري |

(M)

| | |
|-------------------------------------|---------------------------------|
| Machine - aided | الآلي |
| Machine Readable Cataloguing (MARC) | الفهرسة المقروءة آلياً (مارك) |
| Macroevaluation | التقييم الكلي / الشامل |
| Matching | المضاهاة |
| Measures | المقاييس / التدابير |
| Menu | قائمة الاختيار |
| Microcomputers | حاسبات متناهية الصغر |
| Microevaluation | التقييم الجزئي |
| Microfilm | الميكروفيلم |
| Mini Computers | حاسبات مصغرة |
| | الرابط / المحول (ربط الحاسب أو |
| Modem | المنفذ بشبكة الاتصالات) |
| Morphological analysis | التحليل الصرفي |
| Morphological analyzer | المحلل الصرفي |
| Multimedia | الوسائط المتعددة |
| Multipurpose | متعدد الأغراض |

(N)

| | |
|-----------------------------|---|
| National information system | النظام الوطني للمعلومات |
| Natural language | اللغة الطبيعية |
| Near synonyms | أشباه المترادفات |
| Networks | الشبكات |
| Node | النقطة الارتكازية |
| Noise factor | معامل الضوضاء |
| Normalized precision | التحقيق المعياري |
| Normalized recall | الاستدعاء المعياري |
| Novilty ratio | معدل الجودة |
| Number-oriented databases | بنوك المعلومات / مراصد البيانات الرقمية |

(O)

| | |
|--------------|--|
| Obsolescence | التعطّل |
| Offline | خارج الخط المباشر |
| Online | على الخط المباشر |
| OPAC | الفهرس المتاح للجمهور على الخط المباشر |
| Output | المخرجات |

(P)

| | |
|----------------------------|-----------------------------------|
| Patents | براءات الاختراع |
| Performance measures | مقاييس الأداء |
| Pertinency factor | معامل الصلاحية |
| Photoreproduction | الاستنساخ الضوئي |
| Phrase searching | البحث بالعبارات |
| Picture-oriented databases | مراسد بيانات / بنوك معلومات الصور |
| Postcontrolled vocabulary | لغة التقنين اللاحق |
| Postcoordination | الربط اللاحق |
| Precision | التحقيق |
| Precoordination | الربط المسبق |
| Predictability | إمكانية التنبؤ |
| Preprints | الطباعات المبدئية |
| Protocol | قواعد التعامل |
| Prototype | النموذج الأولي |

(Q)

| | |
|----------------------|------------------|
| Qualifier | محدد المعنى |
| Quality | النوعية |
| Quasi - synonyms | أشباه المترادفات |
| Query | الاستفسار |
| Question - answering | الرد على الأسئلة |

(R)

| | |
|-----------------------|----------------------------------|
| Ranked output | المخرجات المرتبة طبقياً |
| Rank recall | الاستدعاء الطبقي |
| Recall | الاستدعاء |
| Recall factor | معامل الاستدعاء |
| Redundancy | الحشو |
| Reference services | الخدمات المرجعية |
| Referring expressions | أدوات تحديد العلاقة |
| Relevance ratio | معدل الصلاحية / الاتصال بالموضوع |
| Resolving power | القدرة التوضيحية |
| Resource sharing | تقاسم الموارد |
| Retrospective search | البحث الراجع |
| Review | المراجعة العلمية |
| Role indicators | محددات / مؤشرات الدور |
| Root | جذع |
| Root words | الكلمات الجذرية |

(S)

| | |
|--|--------------------------------|
| Scatter | التشتت |
| Scope note | تبصرة المجال |
| Screening | التصفية / المراجعة |
| Search logic | منطق البحث |
| Selective dissemination of information (SDI) | البث الانتقائي للمعلومات (بأم) |
| Semantic ambiguity | الغموض الدلالي |
| Semantic road map | خريطة العلاقات الدلالية |
| Simulation | المحاكاة |
| Situational relevance | الصلاحية المشروطة |
| Snobbery ratio | معدل الاقتناع |
| Software | البرمجيات |

| | |
|--------------------|---------------------------|
| Specifity | التخصيص |
| Standardization | التوحيد القياسي / التقييس |
| Statistical phrase | العبارة الإحصائية |
| Stem | جذر الكلمة |
| Stemmer | المجرد (برنامج التجريد) |
| Stemming | التجريد |
| Stop list | قائمة الاستبعاد |
| String | سلسلة / خيط |
| Subheading | رأس فرعي |
| Subject access | الاسترجاع الموضوعي |
| Subject analysis | التحليل الموضوعي |
| Subject heading | رأس الموضوع |
| Subject need | الحاجة الموضوعية |
| Summarization | التلخيص |
| Surrogate | البديل |
| Switching language | لغة التحويل |
| Synonymy | الترادف |
| Syntactic analysis | التحليل النظمي |
| Synthetic | التحليل التركيبي |

(T)

| | |
|--------------------|-----------------------|
| Telecommunications | الاتصالات بعيدة المدى |
| Telefax | التصوير عن بعد |
| Terminal | المنفذ |
| Text indexing | تكشيف النصوص |
| Thesaurus | المكنز |
| Timeliness | الفورية |
| Trade - off | المواءمة |
| Tree structures | التفرعات الشجرية |
| Truncation | البتر |

(U)

Uniterm system

نظام المصطلح الواحد

User - friendly

متعاطف مع المستخدم

User warrant

مسوغ الإفادة

(V)

Vendor

متعهد

Virtual library

المكتبة التصويرية

Vocabulary control

تقنين اللغة

(W)

Weighted indexing

التكشيف الموزون

Word token

هيكل الكلمة

Word type

نوع الكلمة

الكتاب

- يمثل هذا الكتاب امتداداً لكتاب لانكستر الأصلي "نظم استرجاع المعلومات" الذي صدر في طبعتين الأولى عام ١٩٦٨م والثانية عام ١٩٧٩م . وعلى ذلك فإن الكتاب الذي بين أيدينا هو الطبعة الثالثة التي جاءت بعد مضي أربعة عشر عاماً من صدور الطبعة الثانية . و كان للدكتور حشمت قاسم شرف ترجمة هذه الطبعة و التي قبلها . و هي ترصد ما شهدته هذه السنوات من تطورات جوهرية في مراصد البيانات و نظم استرجاع المعلومات . و تقنيات الحاسبات و الاتصالات . و قنوات بث المعلومات بوعي يؤكد تواصل الحلقات . و يبرز تضافر الروافد و المقومات و يمهّد الطريق لما يلي من خطوات .
- وفضلاً عن استيعاب التطورات التي تلت صدور الطبعة الثانية، فإن هذه الطبعة الثالثة تختلف في ميناها و محتواها . فبينما كان الكتاب في طبعته الثانية يضم خمسة و عشرين فصلاً بالإضافة للملاحق الثلاثة ، ينتظم محتوى الطبعة الثالثة اثنا عشر فصلاً . و بينما كان تقييم نظم استرجاع المعلومات موزعاً على ثمانية فصول في الطبعة الثانية ، تركز هذا الموضوع في فصلين اثنين في هذه الطبعة . و بينما كانت لغويات نظم استرجاع المعلومات مشتتة في ثانيا عدد من فصول الطبعة الثانية نجد الاهتمام قد تركز بهذا الموضوع في فصل مستقل في هذه الطبعة . وهذه مجرد أمثلة لمظاهر الاختلاف بين الطبعتين .
- كما حظيت هذه الطبعة بوجود مؤلف مشارك له إسهامانه البارزة في مجال استرجاع المعلومات .